

▼ 9. Nevezetes diszkrét valószínűségi eloszlások

Valószínűségi számítás és statisztika számítógép algebrai támogatással	Készült a PTE PMMIK-n a TÁMOP -4.1.2.A/1-11/1-2011-0098 támogatásával
Klincsik Mihály	

▼ 9.1. Bevezetés

A diszkrét eloszlások közül néhány olyan van, amelyek gyakran előfordulnak az alkalmazások során. A problémák modellezéséhez ki kell választani azt az eloszlást, amelyik a legjobban illeszkedik. Így például gyakran a binomiális és Poisson-eloszlások olyanok, amelyek helyettesíthetők egymással. A feladathoz jobban illeszkedő modellt, akár a gyakoriságok felrajzolásával is eldönthetjük vagy az illesztés jóságának (angolul "goodness of fit") próbájával. (lásd 16. fejezet) A nevezetes diszkrét eloszlások bemutatása mellett tanulmányozzuk alapvető tulajdonságait és alkalmazási területeit. Tárgyaljuk a diszkrét egyenletes, Bernoulli, binomiális, Poisson, hipergeometrikus és geometriai eloszlást. Ezek várható értékét varianciáját meghatározzuk és példákat oldunk meg az eloszlásokkal. A Statistics csomag a diszkrét változókkal végzett műveleteket csak minimális mértékben végzi el.

Célunk az alábbi fogalmak, összefüggések és eljárások megismerése, megértése	
Diszkrét egyenletes eloszlás értelmezése, várható értéke és varianciája	Hipergeometrikus eloszlás és a visszatevés nélküli mintavétel
Bernoulli vagy indikátor eloszlás értelmezése, várható értéke, varianciája.	Poisson eloszlás értelmezése, várható értéke, varianciája
Bernoulli kísérlet alapgondolata	Geometriai eloszlás értelmezése, várható értéke, varianciája
Binomiális eloszlás kétféle származtatása, várható értéke és varianciája	Maple Statistics csomagja segítségével a diszkrét eloszlások hívása, az eloszlások és az eloszlás függvények alkalmazása, rajzolása
Visszatevéses mintavétel és a binomiális eloszlás kapcsolata	Maple Statistics csomag <i>EmpiricalDistribution</i> eljárásának alkalmazása saját diszkrét eloszlások létrehozására

▼ 9.2. A Maple beépített diszkrét valószínűségi eloszlásainak listája és hívása

A valószínűségek számításához Maple-ben a **Statistics** csomag áll rendelkezésre. A csomag betöltése:

> *restart : with(Statistics) :*

A Statistics csomagban található eloszlások helpjét az alábbi módon lehet megtekinteni.

help("Statistics[Distributions]");

A beépített *diszkrét eloszlás típusok* a Maple Statistics csomagban a következők.

--	--	--

Maple hívás neve	elnevezés	paraméterek
DiscreteUniform	egyenletes -eloszlás	a,b
Bernoulli	Bernoulli -eloszlás	p
Binomial	binomiális eloszlás	n,p
Geometric	geometriai eloszlás	p
Hypergeometric	hipergeometrikus eloszlás	N,s,n
Poisson	poisson - eloszlás	λ
NegativeBinomial	negatív binomiális	r,p

A diszkrét valószínűségi eloszlásokat *Distribution(Maple név(paraméterek))* formában lehet hívni. Nézzük a binomiális eloszlás alkalmazását példaként.

9.2.1. PÉLDA. Binomiális eloszlás megadása

- (a) Adjuk meg a binomiális eloszlást $n = 7$ és $p = 0.2$ paraméterekkel! Milyen adatstruktúra keletkezett?
(b) Rajzoljuk fel az **eloszlás** pálcika diagramját és az **eloszlás függvény** lépcsős diagramját!

Megoldás

(a) Adjuk meg a binomiális eloszlást $n = 7$ és $p = 0.2$ paraméterekkel! Milyen adatstruktúra keletkezett?

> *B := Distribution(Binomial(7, 0.2));*

B := module() (1.2.1)

option Distribution, Discrete;

export Conditions, ParentName, Parameters, CDF, CharacteristicFunction, Kurtosis, Mean, MGF, ProbabilityFunction, Skewness, Support, Variance, VariationCoefficient, CDFNumeric, QuantileNumeric, RandomSample, RandomSampleSetup, RandomVariate;

end module

Látható, hogy B egy modul, amely a binomiális eloszláshoz készített eljárások gyűjteményét tartalmazza. Ebből valamelyik eljárást *B:-eljárás név* alakban lehet meghívni. Például a *B:-Parameters* eljárás visszaadja a B eloszlás paramétereit. A *B:-Conditions* eljárás a B eloszlás paramétereire kirótt feltételt kírja és ez alapján ellenőrizni lehet, hogy az általunk megadott paraméterek teljesítik-e a feltételeket!

> *B:-Parameters, B:-Conditions*

[7, 0.2000000000], [0 < 7, 0 ≤ 0.2000000000, 0.2000000000 ≤ 1, 7::integer] (1.2.2)

Látható, hogy az $n=7$ és $p=0.2$ paramétereket adtuk meg. Az $n > 0$ és egész szám, valamint $0 \leq p \leq 1$ feltételek jelen esetben teljesülnek.

(b) Rajzoljuk fel az eloszlás pálcika diagramját és az eloszlás függvény lépcsős diagramját!

A "ProbabilityFunction" eljárás megadja képletek formában az **eloszlást**. Az alábbi két hívási szekvenciát alkalmazhatjuk. Mindkét hívás eredménye ugyanaz.

> *ProbabilityFunction(B, x)*

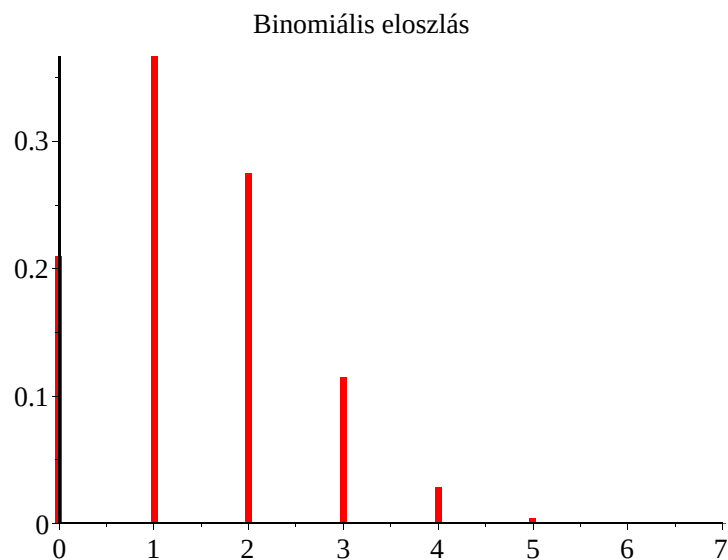
$$\begin{cases} 0 & x < 0 \\ \text{binomial}(7, x) 0.2^x 0.8^{7-x} & \text{otherwise} \end{cases} \quad (1.2.3)$$

> *f := B:-ProbabilityFunction(x);*

$$f := \begin{cases} 0 & x < 0 \\ \text{binomial}(7, x) 0.2^x 0.8^{7-x} & \text{otherwise} \end{cases} \quad (1.2.4)$$

Az eloszlás pálcika diagramja az *f* diszkrét függvény képlete alapján rajzolható!

plot([seq([x, 0], [x, f]), x = 0 .. 7), thickness = 3, color = red, title = "Binomiális eloszlás")

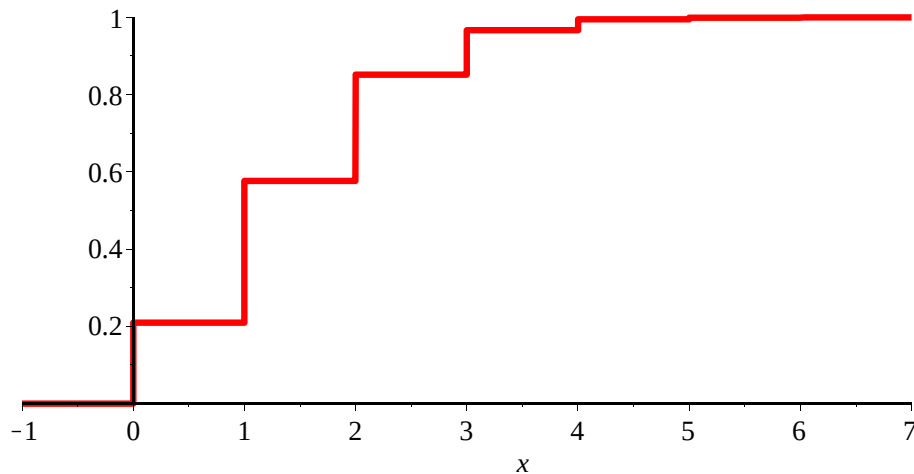


Az **eloszlásfüggvényt** a CDF(B,x) eljárással kapjuk, amelyből a lépcsős függvény rajzolható!

> *F := CDF(B, x)*

$$F := \begin{cases} 0 & x < 0 \\ 1 & 7 \leq x \\ \sum_{i=0}^{\text{floor}(x)} \text{binomial}(7, i) 0.2000000000^i 0.8000000000^{7-i} & \text{otherwise} \end{cases} \quad (1.2.5)$$

> *plot(F, x = -1 .. 7, 0 .. 1, thickness = 3, color = red);*



9.2.2. PÉLDA. Saját készítésű diszkrét eloszlás

Készítsünk olyan S diszkrét eloszlást, amely az

$$X = [1, 2.5, 4, 5]$$

helyeken rendre felveszi az alábbi valószínűségeket

$$P = \left[\frac{1}{3}, \frac{1}{4}, \frac{13}{60}, \frac{1}{5} \right].$$

Tehát $P(S = X_i) = P_i$ kell, hogy teljesüljön, ahol $i = 1, 2, 3, 4$.

(a) Számoljuk ki S várható értékét!

(b) Az S eloszlásból nyerjük ki X és P értékeit!

(c) Rajzoljuk fel az S eloszlás túske diagramját!

(d) Rajzoljuk fel S eloszlás függvényének lépcsős diagramját!

Megoldás

A *Statistics* csomagban található az *EmpiricalDistribution* eljárás segítségével tudunk megadni egy S saját eloszlást!

> *restart : with(Statistics) :*

> $S := \text{Distribution}(\text{EmpiricalDistribution}(\langle 1, 2.5, 4, 5 \rangle, 'probabilities' = [1/3, 1/4, 13/60, 1/5]))$

$S := \text{module}(\)$ (1.2.6)

option *Distribution, Discrete;*

export *Conditions, ParentName, Parameters, CDF, DiscreteValueMap, Mean,*

Median, Mode, ProbabilityFunction, Quantile, Support, RandomSample,

RandomVariate;

end module

(a) Most is kaptunk egy modult és benne a hívható eljárások! Nézzük a várható érték számítását a *Mean* eljárás megadja-e automatikusan!

> $\text{átlag} := \text{Mean}(S)$

$\text{átlag} := 2.825000000000000$

(1.2.7)

(b) Az S modulból az X értékeket a *DiscreteValueMap* eljárással kaphatjuk meg.

> *E* := *DiscreteValueMap*(*S*, 'n');

> *X* := [seq(eval(*E*, *n* = *k*), *k* = 1 ..4)]

$$E := \left[1, \frac{5}{2}, 4, 5 \right]_n$$

$$X := \left[1, \frac{5}{2}, 4, 5 \right] \quad (1.2.8)$$

A *ProbabilityFunction* eljárás a valószínűségeket megadja.

> *P* := [seq(*ProbabilityFunction*(*S*, *X*_{*i*}), *i* = 1 ..4)]

$$P := \left[\frac{1}{3}, \frac{1}{4}, \frac{13}{60}, \frac{1}{5} \right] \quad (1.2.9)$$

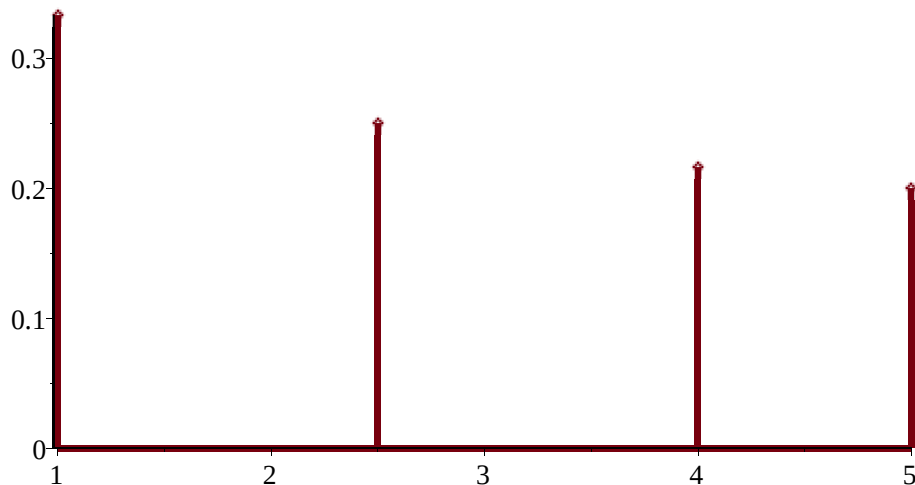
> Sum('P', *i* = 1 ..numelems(*P*)) = add(*P*_{*i*}, *i* = 1 ..numelems(*P*))

$$\sum_{i=1}^4 P_i = 1 \quad (1.2.10)$$

(c) Rajzoljuk fel az S eloszlás túske diagramját!

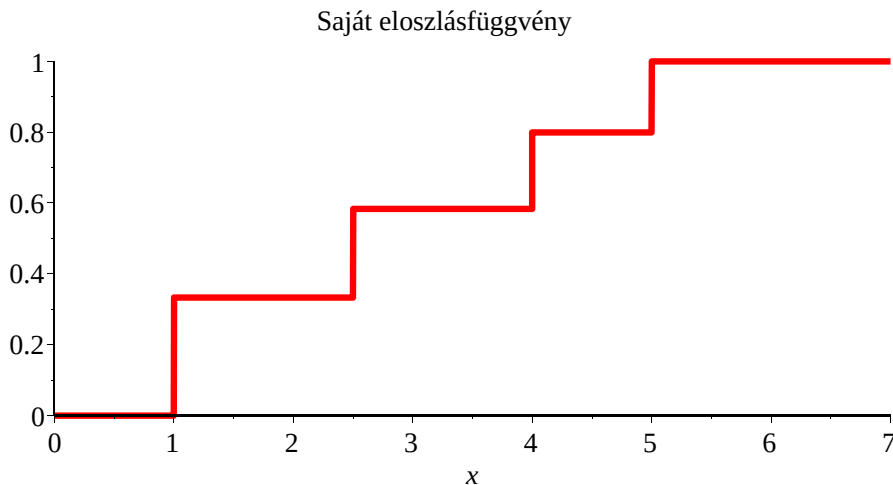
Az S diszkrét eloszlás X értékei és a hozzá tartozó P valószínűségek ismeretében a pálcika vagy túske diagram a *DiscretePlot* eljárással megkapható, mely a *DynamicSystems* csomagban van!

> *DynamicSystems:-DiscretePlot*(*X*, *P*, style = stem, thickness = 3)



A saját modul CDF eljárása segítségével a lépcsősfüggvény felrajzolható!

> plot(CDF(*S*, *x*), *x* = 0 ..7, thickness = 3, color = red, title = "Saját eloszlásfüggvény")



▼ 9.3. Diszkrét egyenletes eloszlás

Tegyük fel, hogy az X valószínűségi változó lehetséges értékei $x_1, x_2, \dots, x_n$ és ezeket azonos

$p_1 = p_2 = \dots = p_n$ valószínűséggel veszi fel. Mivel $\sum_{k=1}^n p_k = 1$, ezért mindegyik p_k érték $\frac{1}{n}$ kell, hogy legyen.

9.3.1. DEFINÍCIÓ. Egyenletes eloszlás

Az alábbi táblázattal adott

$$X = \begin{bmatrix} X & x_1 & x_2 & \dots & x_k & \dots & x_n \\ P & \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix}$$

diszkrét eloszlást diszkrét egyenletes eloszlásnak nevezzük, ahol

$$P(X = x_1) = P(X = x_2) = \dots = P(X = x_n) = \frac{1}{n}.$$

Így egy diszkrét egyenletes eloszlású valószínűségi változó megadásához szükséges paraméterek: az $x_1, x_2, \dots, x_n$ értékek!

Teljes általánosságában a Maple-ben csak saját eloszlásként tud ilyet előállítani. Viszont olyan speciális eloszlást, amelyre $x_i = a + (i - 1)$ egymásután következő egész számok tudunk készíteni. Például a szabályos kocka dobás szimulációját megadjuk.

9.3.2. PÉLDA. Szabályos kocka dobás

Tekintsük a szabályos kocka dobás kísérletét! Legyen X a dobott pontszám értéke.

X lehetséges kimenetele: 1, 2, 3, 4, 5, 6 és ezeket egyforma $\frac{1}{6}$ valószínűséggel veszi fel.

- (a) Adjuk meg az X valószínűségi változót!
- (b) Rajzoljuk fel X eloszlását és eloszlás függvényét!
- (c) Számoljuk ki X várható értékét és szórását!

Megoldás

(a) A `DiscreteUniform(a, b)` az egyenletes eloszlás hívása.

> `restart : with(Statistics) :`

`X := RandomVariable(DiscreteUniform(1, 6)) ;`

`X := _R`

(1.3.1)

A `ProbabilityFunction` eljárás valójában minden 1 és 6 közötti x értékre $\frac{1}{6}$ értéket vesz fel folytonosan!

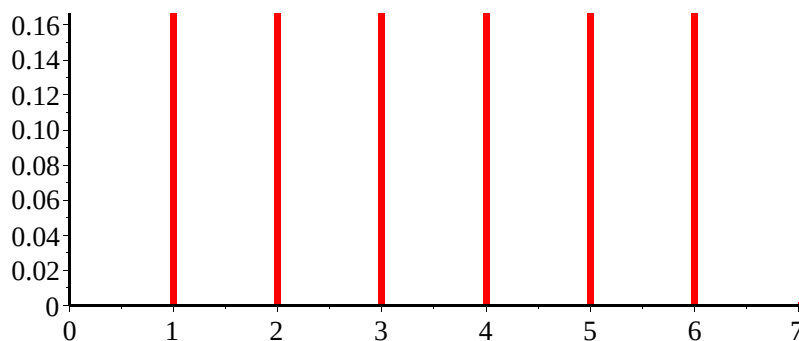
> `p := ProbabilityFunction(X, x);`

$$p := \begin{cases} 0 & x < 1 \\ \frac{1}{6} & x \leq 6 \\ 0 & 6 < x \end{cases}$$

(1.3.2)

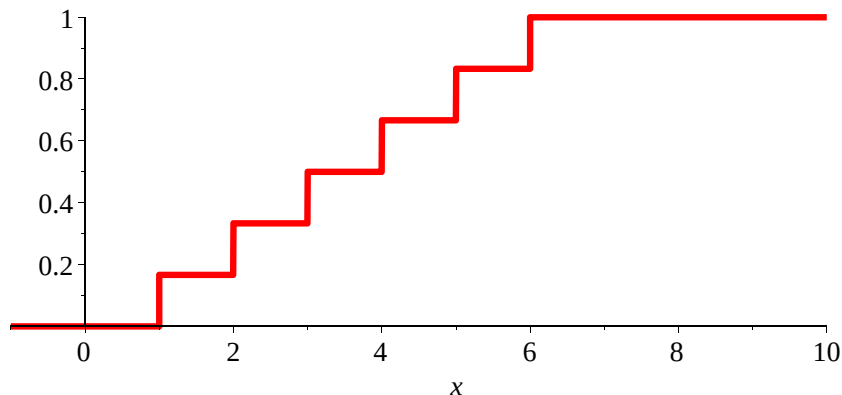
(b) Az eloszlás pálcika diagramjához el kell készítenünk a diszkrét érték készletet! A valószínűségek a p kifejezésben vannak!

> `plot([seq([x, 0], [x, p]), x = 0..7], thickness = 3, color = red);`



Az eloszlás függvény egyenletes lépésközökkel emelkedő lépcső.

> `plot(CDF(X, x), x = -1..10, 0..1, thickness = 3, color = red);`



(c) Számoljuk ki X várható értékét és szórását!

> várható érték, szórás = $Mean(X)$, $StandardDeviation(X)$

$$\text{várható érték, szórás} = \frac{7}{2}, \frac{1}{6} \sqrt{105} \quad (1.3.3)$$

9.3.3. TÉTEL. Egyenletes eloszlás várható értéke és varianciája

Igazoljuk, hogy az $X = [x_1, x_2, \dots, x_n]$ pontokban egyenletes eloszlás várható értéke

$$E(X) = \frac{x_1 + x_2 + \dots + x_n}{n}$$

átlag és a varianciája

$$Var(X) = \frac{\sum_{k=1}^n x_k^2}{n} - \left(\frac{\sum_{k=1}^n x_k}{n} \right)^2.$$

$$\text{Ha } x_i = i, \text{ akkor } E(X) = \frac{n+1}{2} \text{ és } Var(X) = \frac{n^2-1}{12}.$$

Bizonyítás

A diszkrét eloszlások várható érték számítási szabálya alapján

$$\mu = E(X) = \sum_{i=1}^n x_i \cdot p_i = \sum_{i=1}^n x_i \cdot \frac{1}{n} = \frac{\sum_{i=1}^n x_i}{n}.$$

Ha $x_i = i$, akkor

$$\mu = E(X) = \frac{1 + 2 + \dots + n}{n} = \frac{\frac{n \cdot (n+1)}{2}}{n} = \frac{n+1}{2}.$$

A variancia számítási képletét alkalmazzuk

$$Var(X) = E(X^2) - E(X)^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{k=1}^n x_k}{n} \right)^2.$$

Ha $x_i = i$, akkor

$$Var(X) = \frac{\sum_{i=1}^n i^2}{n} - \left(\frac{n+1}{2} \right)^2 = \frac{\frac{1}{6} n (n+1) (2n+1)}{n} - \frac{(n+1)^2}{4} =$$

$$\frac{1}{6} \cdot (n+1) \cdot (2n+1) - \frac{(n+1)^2}{4} = \frac{(n+1)}{2} \cdot \left(\frac{2n+1}{3} - \frac{n+1}{2} \right) = \frac{1}{12} (n+1) \cdot (n-1)$$

Q.e.d.

▼ 9.4. Bernoulli eloszlás

Tekintsünk egy kísérletet, melynek két lehetséges kimenetele van, **sikeres** vagy **sikertelen**. A kísérlet sikeres kimenetelének a bekövetkezési valószínűsége legyen p ($0 < p < 1$). Ekkor a sikertelen kimenetel bekövetkezési valószínűsége $(1-p) = q$.

Az ilyen X valószínűségi változót Bernoulli-típusúnak nevezzük és X eloszlása $P(X=1) = p$ és

$$P(X=0) = 1 - p = q.$$

9.4.1 DEFINÍCIÓ. Bernoulli-változó

Az X valószínűségi változót Bernoulli-típusúnak nevezzük $p \in (0, 1)$ paraméterrel, ha az $X = 1$ és $X = 0$ értékeket veszi fel és $P(X = 1) = p$, $P(X = 0) = 1 - p = q$. Tehát X eloszlása mátrix alakban

$$\begin{bmatrix} X & 1 & 0 \\ P & p & 1 - p \end{bmatrix}$$

Az $X \sim \text{Bernoulli}(p)$ jelölést használjuk arra, hogy az X változó eloszlása Bernoulli-típusú p -paraméterrel.

Maple-ben az X **Bernoulli-típusú** valószínűségi változót a következőképpen tudjuk felvenni!

> *restart* : *with(Statistics)* :

> $X := \text{RandomVariable}(\text{Bernoulli}(p))$

$X := _R$

(1.4.1)

Az egyes értékek valószínűségeit a **ProbabilityFunction** adja meg!

> $f := \text{ProbabilityFunction}(X, x);$

$P('X'=1) = \text{ProbabilityFunction}(X, 1),$

$P('X'=0) = \text{ProbabilityFunction}(X, 0);$

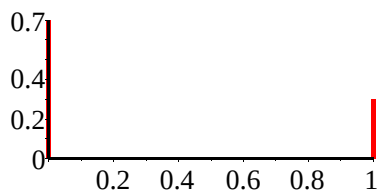
$$f := \begin{cases} 0 & x < 0 \\ 0 & 1 < x \\ (2p - 1)x + 1 - p & \text{otherwise} \end{cases}$$

$$P(X=1) = p, P(X=0) = 1 - p$$

(1.4.2)

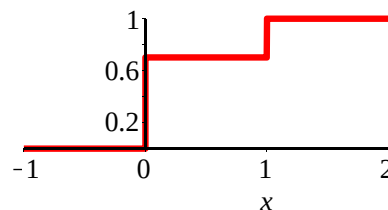
Az eloszlás pálcika diagramja

> $\text{plot}(\text{subs}(p = 0.3, [\text{seq}([x, 0], [x, f]), x = 0..1]), \text{thickness} = 2, \text{color} = \text{red})$



Az eloszlásfüggvény lépcsős függvénye

> $\text{plot}(\text{CDF}(\text{Bernoulli}(0.3), x), x = -1..2, \text{thickness} = 3, \text{color} = \text{red})$



A várható értéket és a varianciát a diszkrét eloszlás képletei alapján egyszerűen kapjuk.

9.4.2. TÉTEL. Bernoulli-változó várható értéke és varianciája

Ha $X \sim \text{Bernoulli}(p)$, akkor

$$E(X) = p \text{ és } \text{Var}(X) = p \cdot (1 - p)$$

várható értéke és varianciája.

Bizonyítás.

A várható értéket az $E(X) = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$ képletekkel számolhatjuk ki. Ugyanezt

adja a beépített **Mean** eljárás is!

> *Várható érték* = *Mean*(*X*);

$$\text{Várható érték} = p \quad (1.4.3)$$

A variancia vagy szórásnégyzet számításához a következő képletek vezetnek

$$\text{Var}(X) = E(X^2) - E(X)^2 = p - p^2 = p(1 - p),$$

mert X és X^2 eloszlása megegyezik.

> *Variancia* = *Variance*(*X*);

Szórás = *StandardDeviation*(*X*);

$$\text{Variancia} = p(1 - p)$$

$$\text{Szórás} = \sqrt{p(1 - p)} \quad (1.4.4)$$

Q.e.d.

▼ 9.5. Binomiális eloszlás

Az egyik leggyakrabban előforduló speciális diszkrét valószínűségi eloszlás a binomiális eloszlás.

Binomiális eloszlást kapunk, ha egy **A** véletlen eseményre "**n**" **független** kísérletet végzünk és azt szeretnénk tudni, hogy hányszor következik be az **A** esemény és egyetlen bekövetkezés valószínűsége p , azaz $P(A) = p$.

9.5.1. DEFINÍCIÓ. BERNOULLI-FÉLE KÍSÉRLET

(i) Figyeljük egy véletlen kísérlet valamely A eseményének bekövetkezését. Az A esemény bekövetkezési valószínűsége legyen $p = P(A)$.

(ii) A kísérletet megismételjük egymástól függetlenül n -szer egymásután.

A Bernoulli-kísérlettel kapcsolatban többféle valószínűségi változót is definiálhatunk. Ha azt figyeljük, hogy az n kísérletből hányszor következik be az A esemény, akkor kapunk binomiális eloszlást. Ha azt figyeljük, hogy az A esemény hányadikra következik be először, akkor geometriai eloszlást kapunk. Ezért a Bernoulli-eloszlásba nem értjük bele azt a feltételt, amely a binomiális eloszlást értelmezi.

Jelölje $q = 1 - p = P(\bar{A})$ annak valószínűségét, hogy A nem következik be.

9.5.2. DEFINÍCIÓ. Binomiális eloszlás n és p paraméterekkel

Tekintsük a Bernoulli-kísérlet során végzett n független kísérletből azon kísérletek X számát, ahányszor bekövetkezett az A esemény. Tehát az X valószínűségi változó értéke az A bekövetkezési száma az n független kísérletből,

X = az " A " esemény bekövetkezési száma n független kísérletből.

Így X értéke $0, 1, 2, 3, \dots, n$ lehet.

Ebben az esetben az $X \sim \text{binomial}(n, p)$ jelölést használjuk.

Határozzuk meg az egyes $X = k$ értékekhez tartozó valószínűségeket! Az egyszerűség kedvéért $B = \bar{A}$ jelölje az A ellentét eseményét! Ekkor az

$$\{X = k\}$$

esemény, vagyis az A esemény k -szor következik be úgy, hogy az $\underbrace{A, A, \dots, A, B, B, \dots, B}_{n \text{ darab}}$ " n "

hosszúságú sorozatokban pontosan k -darab A betű és $(n-k)$ darab B betű van.

A k -darab " A " betű helyét $\binom{n}{k}$ -féleképpen választhatjuk ki. Egy kiválasztás valószínűsége a

függetlenség miatt

$$P(A A \dots A B B \dots B) = P(A) \cdot P(A) \cdot \dots \cdot P(A) \cdot P(B) \cdot P(B) \dots \cdot P(B) = p^k q^{n-k}$$

Mivel az ilyen kiválasztások egymást kizárók, ezért az $\binom{n}{k}$ darab valószínűséget össze kell adni, vagyis

$$P(X=k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}.$$

Tehát X eloszlását felírhatjuk a következő táblázatos formában

$$\begin{bmatrix} X & 0 & 1 & \dots & k & \dots & n \\ P & \binom{n}{0} p^0 q^n & \binom{n}{1} p q^{n-1} & \dots & \binom{n}{k} p^k q^{n-k} & \dots & \binom{n}{n} p^n \end{bmatrix}$$

ahol $\binom{n}{k}$ jelöli az n alatt a k binomiális együtthatót.

9.5.3. TÉTEL. A binomiális eloszlás értékeinek számítása

Ha az $X \sim \text{binomial}(n, p)$, vagyis az X valószínűségi változó (n, p) paraméterű **binomiális eloszlású diszkrét valószínűségi változó**, akkor

$$P(X=k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}, \text{ ahol } k=0, 1, 2, \dots, n \text{ lehet.}$$

Képezzük ezt a formulát szimbolikus módon, adott n, k és $0 < p < 1$ esetén!

> *restart; with(plots) : with(combinat) : with(Statistics) :*

$$\text{Binomial_Formula} := (n, p, k) \rightarrow \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k};$$

$$\text{Binomial_Formula} := (n, p, k) \rightarrow \text{combinat:-binomial}(n, k) p^k (1-p)^{n-k} \quad (1.5.1)$$

Ezzel a formulával valóban diszkrét eloszlást adtunk meg, mert összegezve a valószínűségeket 1-t kapunk, melyet az $(a+b)^n$ két tag összegnek hatványáról szóló binomiális -tétel alapján kapunk

$$(p + 1 - p)^n = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = 1$$

a = p és b=1-p választással.

> *assume(p < 1) :*

$$\begin{aligned} > \sum_{k=0}^n \text{Binomial_Formula}(n, p, k) = \sum_{k=0}^n \text{Binomial_Formula}(n, p, k); \\ & \text{lhs}(\%) = \text{simplify}(\text{rhs}(\%)) \end{aligned}$$

$$\sum_{k=0}^n \text{binomial}(n, k) p^k (1-p)^{n-k} = \left(\frac{p}{1-p} + 1 \right)^n (1-p)^n$$

$$\sum_{k=0}^n \text{binomial}(n, k) p^k (1-p)^{n-k} = 1 \quad (1.5.2)$$

Az eloszlást meghatározó képleteket megkapjuk a **Statistics** csomag **"ProbabilityFunction"** eljárásával is!

> *X := RandomVariable(Binomial(n, p)) :*

> *ProbabilityFunction(X, k)*

$$\begin{cases} 0 & k < 0 \\ \text{binomial}(n, k) p^k (1-p)^{n-k} & \text{otherwise} \end{cases} \quad (1.5.3)$$

9.5.4. TÉTEL. Független azonos paraméterű Bernoulli-eloszlások összege binomiális-eloszlás

Ha X_k jelöli a k -ik független kísérlet kimenetelét, akkor $X_k \sim \text{Bernoulli}(p)$ minden $k = 1, 2, \dots, n$ esetén. Továbbá az

$$X = X_1 + X_2 + \dots + X_n$$

összege adott változó binomiális n és p paraméterekkel.

Bizonyítás

A Bernoulli kísérlet során k -ik eredmény

$$X_k = \begin{cases} 1, & \text{ha az A esemény bekövetkezett a } k\text{-ik kísérletre} \\ 0, & \text{ha az A esemény nem következett be a } k\text{-ik kísérletre} \end{cases}$$

egy Bernoulli-változó.

Az $X = X_1 + X_2 + \dots + X_n$ összegben ott van 1, ahol A bekövetkezett és 0 ott, ahol nem következett be az A esemény. Ezért X értéke az n kísérletből az A bekövetkezési száma, azaz binomiális eloszlású n és p paraméterekkel. Q.e.d.

9.5.5. TÉTEL. A binomiális-eloszlás várható értéke és varianciája

Az $X \sim \text{binomial}(n, p)$ változó várható értéke és varianciája

$$E(X) = n \cdot p \text{ és } \text{Var}(X) = n \cdot p \cdot (1 - p).$$

Bizonyítás

Tekintsük az $X \sim \text{binomial}(n, p)$ változó

$$X = X_1 + X_2 + \dots + X_n$$

felbontását $X_k \sim \text{Bernoulli}(p)$ változók összegére. Mivel a várható érték additív függvény, ezért

$$E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = n \cdot p,$$

mert a Bernoulli változókra $E(X_k) = p$ teljesül.

Maple összegzéssel is kiszámíthatjuk, hogy az n és p paraméterű binomiális eloszlás **várható értéke** = $\mu = n p$ vagy a "Mean" eljárással közvetlenül is megkaphatjuk

> *Sum(k·Binomial_Formula(n, p, k), k = 0..n) = sum(k·Binomial_Formula(n, p, k), k = 0..n);*
Várható_érték := simplify(rhs(%));
Ugyanez a Mean eljárással = Mean(X);

$$\sum_{k=0}^n k \text{binomial}(n, k) p^k (1-p)^{n-k} = \frac{\left(\frac{p}{1-p} + 1\right)^n p n (1-p)^n}{(1-p) \left(\frac{p}{1-p} + 1\right)}$$

$$\text{Várható_érték} := n p$$

$$\text{Ugyanez a Mean eljárással} = n p$$

(1.5.4)

Megmutatjuk, hogy az n és p paraméterű binomiális eloszlás **szórásnégyzete** $\sigma^2 = n p (1 - p)$.

Mivel a Bernoulli-kísérletben függetlenek a kísérletek, ezért az $X_1, X_2, \dots, X_n$ valószínűségi változók is páronként függetlenek, ezért a varianciát tagonként vehetjük

$$\begin{aligned}\text{Var}(X) &= \text{Var}(X_1 + X_2 + \dots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) = \\ &= p(1-p) + p(1-p) + \dots + p(1-p) = np(1-p).\end{aligned}$$

Ugyanez Maple-ben.

> $\text{Sum}(k^2 \cdot \text{Binomial_Formula}(n, p, k), k = 0 .. n) = \text{simplify}(\text{sum}(k^2 \cdot \text{Binomial_Formula}(n, p, k), k = 0 .. n))$;
Szórás_négyzet = $\text{factor}(\text{rhs}(\%) - \text{Várható_érték}^2)$;

$$\sum_{k=0}^n k^2 \text{binomial}(n, k) p^k (1-p)^{n-k} = np + p^2 n^2 - p^2 n$$

$$\text{Szórás_négyzet} = -np(-1 + p) \quad (1.5.5)$$

Ugyanezt adja a "Variance" eljárás

> *A variancia a Variance eljárással* = $\text{Variance}(X)$
A variancia a Variance eljárással = $np(1-p)$ (1.5.6)

Q.e.d.

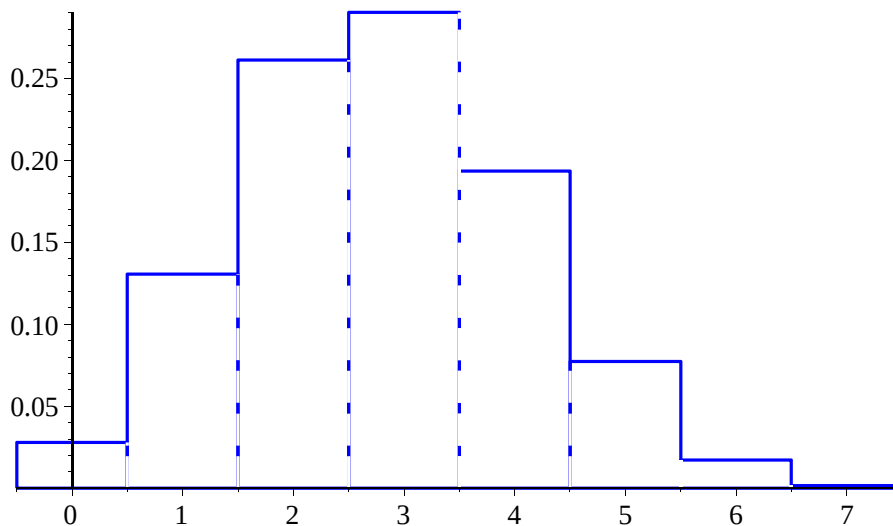
9.5.6. HISZTOGRAM, ELOSZLÁS

Binomiális eloszlás hisztogramjának ábrázolására készítettünk egy eljárást, melynek paraméterei az n és p .

```
> BiHist := proc(n, p)
  local f, k, c, d, h, P, Q;
  c := blue; P := NULL : Q := NULL :
  for k from 0 to n do
    h := binomial(n, k) * (p^k) * ((1-p)^(n-k));
    P := P, plot( [[k-.5, 0], [k+.5, 0], [k+.5, h], [k-.5, h], [k-.5, 0]], color = c );
    if (k < n/2)
      then Q := Q, plot( [[k+.5, 0], [k+.5, h]],
        color = white, linestyle = 3) :
      else Q := Q, plot( [[k-.5, 0], [k-.5, h]],
        color = white, linestyle = 3) : fi;
  od;
  plots[display]( [ P, Q ] );
end proc;
```

A hisztogramot nézzük $p = 0.4$ és $n = 7$ esetére!

> $\text{BiHist}(7, 0.4)$



9.5.7. HISZTOGRAM ANIMÁCIÓK VÁLTOZÓ PARAMÉTEREK MELLETT

Vizsgáljuk az eloszlás alakját és formáját
állandó $p = 0.4$ valószínűség és növekvő n
esetén $n = 4$ -től $n = 20$ között!

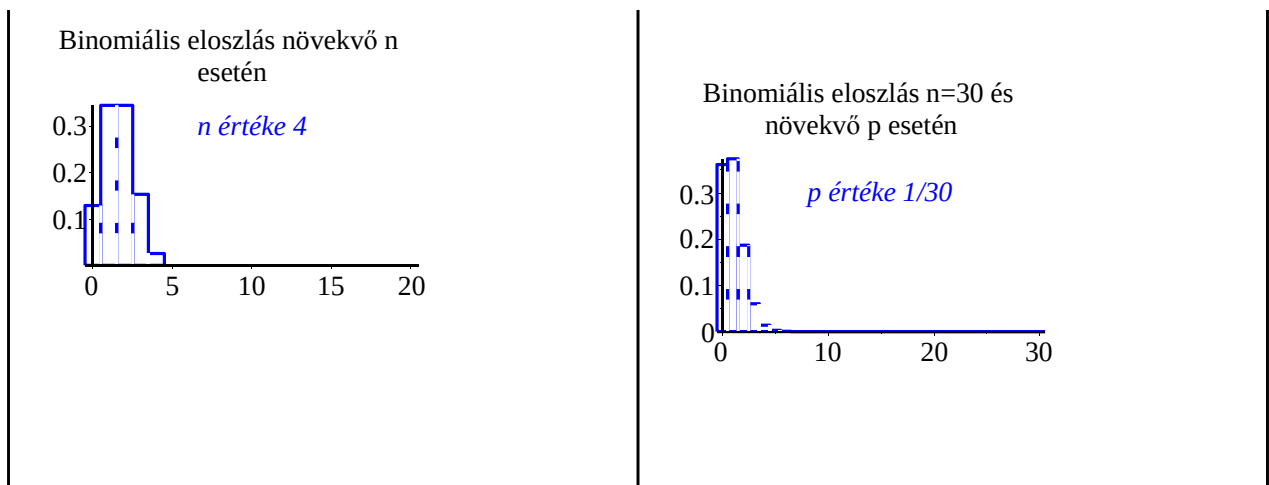
```
> p := 0.4; N := 20 :
> for n from 4 to N do
>   Hn := BiHist(n, 0.4) :
>   num := convert(n, string) :
>   trackern := textplot( [ [  $\frac{N}{2}$ , 0.3,
                             cat(n értéke, num) ], color = blue ) :
>   Pn := display( { Hn, trackern } ) :
> od :
> display( [ seq( Pn, n = 4 .. N ) ],
           insequence = true, title
           = "Binomiális eloszlás növekvő n
           esetén")
```

$p := 0.4$

Vizsgáljuk az eloszlás alakját és formáját
állandó $n = 30$ valószínűség és növekvő p
esetén $p = 1/30$ -től $p = 1/2$ -ig $1/30$ -ad
lépésközökkel!

```
> leptek :=  $\frac{1}{30}$ ; n := 30 :
> for i from 1 to 15 do
>   prob := i * leptek :
>   H[i] := BiHist(n, prob) :
>   probs := convert(prob, string) :
>   tracker[i] := textplot( [ 15, 0.3,
                              cat(p értéke, probs) ], color = blue)
>   :
>   P[i] := display( { H[i], tracker[i] } ) :
> od :
> display( [ seq( P[i], i = 1 .. 15 ) ],
           insequence = true, title
           = "Binomiális eloszlás n=30 és
           növekvő p esetén")
```

$leptek := \frac{1}{30}$



9.5.8. A BINOMIÁLIS ELOSZLÁS KÖZELÍTÉSE NORMÁL ELOSZLÁSSAL

Fix $p = 0.2$ valószínűséggel növekvő n mellett felrajzoljuk a $\text{binomial}(n, p)$ binomiális eloszlás hisztogramját és az ugyanolyan várható értékű és szórású normál eloszlás sűrűség függvényét!

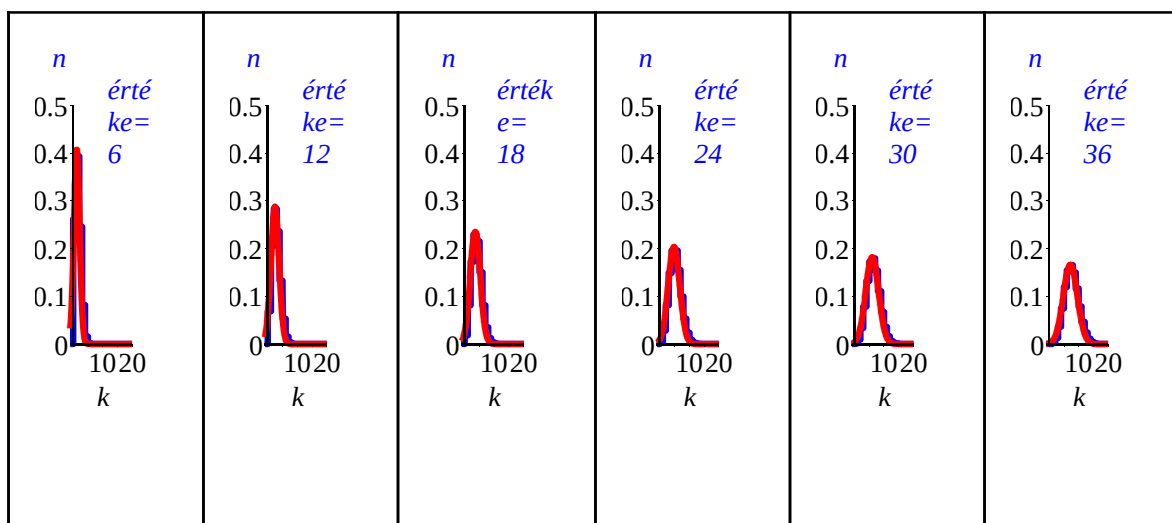
- ```

> restart : with(Statistics) : with(plots) :
> p := 0.2
> kepek := NULL :
 for n from 6 to 36 by 6 do
> kep1 := plot(ProbabilityFunction(Binomial(n, p), floor(k)), k = -1 .. 20, color = blue,
 thickness = 3) :
> kep2 := plot(PDF(Normal(n·p, sqrt(n·p·(1 - p))), x), x = -1 .. 20, color = red, thickness
 = 3) :

> num := convert(n, string) : szöveg := textplot([10, 0.5, cat(n értéke=, num)], color
 = blue) :
> kepek := kepek, display([kep1, kep2, szöveg]) :
 end do:
> display(Matrix(1, 6, [kepek]))

```

$p := 0.2$



Valóban  $n$  növekedésével - fix  $p$  valószínűség érték mellett - a  $\text{Binomial}(n, p)$  eloszlást egyre

jobban megközelíti a  $\text{Normal}(\mu = np, \sigma = \sqrt{np(1-p)})$  normál eloszlás ugyanolyan várható értékkel és szórással, mint a binomiális eloszlás. Az ábra szemléletes bizonyítást ad a centrális határeloszlás-tételre.

### 9.5.9. MINTAVÉTELEZÉS VISSZATEVÉSSEL

A **visszatevéses mintavételezés** során is binomiális eloszlást kapunk.

Tekintsünk egy dobozt, melyben  $N$  darab alkatrész pl. csavar van.

A dobozban levő alkatrészek között  $S$  darab alkatrész ( $0 < S \leq N$ ) megkülönböztetett tulajdonságú, például **selejtes**.

Egy selejtes alkatrész húzásának valószínűsége  $p = \frac{S}{N}$ .

Húzzunk  $n$  alkalommal egy csavart véletlenszerűen úgy, hogy minden húzás után visszatesszük a kihúzott csavart. Ekkor legyen az  $A$  esemény a selejtes csavar húzásának eseménye. Mivel  $P(A)$

$= p = \frac{S}{N}$ , ezért annak valószínűsége, hogy az  $n$  kihúzott csavar között  $k$  selejtes van binomiális eloszlást ad

$$P(n \text{ kihúzott csavar között } k \text{ darab selejtes van}) = \binom{n}{k} \cdot \left(\frac{S}{N}\right)^k \cdot \left(1 - \frac{S}{N}\right)^{n-k} \quad (k = 0, 1, 2, \dots, n).$$

### 9.5.10. PÉLDA. Működés valószínűsége

Egy új ház tulajdonos 20 lámpában cserélt izzót egyidejűleg. Tegyük fel, hogy mindegyik izzó  $p = 0.2$  valószínűséggel működik 3 hónapnál tovább.

(a) Mekkora a valószínűsége, hogy a 20 izzó közül legalább 5 izzó működik 3 hónappal később?

(b) Átlagosan hány izzót kell a 3 hónap alatt kicserélnie?

### MEGOLDÁS

Az  $X$  valószínűségi változót válasszuk meg úgy, hogy illeszkedjen az (a) pontbeli kérdéshez! Jelölje  $X$  azon izzók számát, amelyek a 20 izzó közül 3 hónapnál **tovább működnek**. Mivel az izzók működési ideje előre nem kiszámítható, ezért  $X$  véletlen valószínűségi változó. Ha  $A$  jelöli azt az eseményt, hogy egy izzó 3 hónap eltelte után is működik, akkor az  $A$  eseményre  $n=20$  független kísérletet végzünk és figyeljük, hogy hányszor következik be az  $A$  esemény. Ezért  $X$  eloszlása binomiális  $n=20$  és  $p=0.2$  paraméterekkel. Az (a) kérdés a  $P(5 \leq X)$  valószínűséget kérdezi, mely a kiegészítő esemény segítségével ( $1 - P(X < 5)$ ) az eloszlás függvényből számolható (1-CDF(4)) képlettel.

> *restart : with(Statistics) :*

>  $n, p := 20, 0.2$

$$n, p := 20, 0.2 \quad (1.5.7)$$

>  $X := \text{Binomial}(n, p) :$

>  $f := \text{unapply}(\text{ProbabilityFunction}(X, u), u) : 'f(u) = f(u)$

$$f(u) = \begin{cases} 0 & u < 0 \\ \text{binomial}(20, u) 0.2^u 0.8^{20-u} & \text{otherwise} \end{cases} \quad (1.5.8)$$

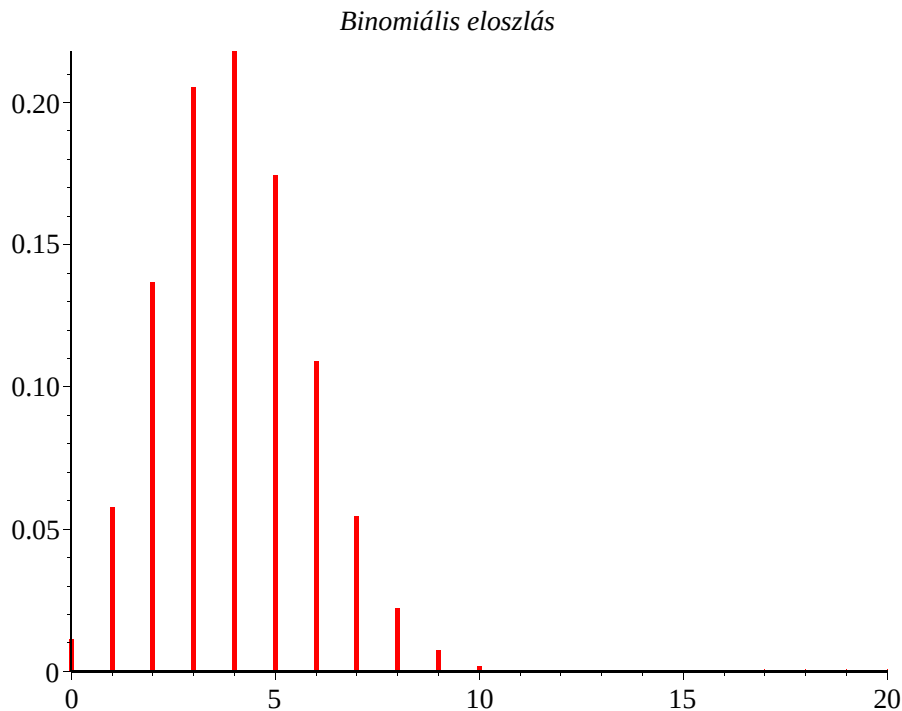
>  $px := [\text{seq}(f(k), k = 0..20)]$

$$px := [0.01152921505, 0.05764607524, 0.1369094287, 0.2053641430, 0.2181994019, 0.1745595216, 0.1090997010, 0.05454985049, 0.02216087676, 0.007386958920, 0.002031413703, 0.0004616849325, 0.00008656592484, 0.00001331783459, \dots] \quad (1.5.9)$$



0.000001664729324,  $1.664729324 \cdot 10^{-7}$ ,  $1.300569784 \cdot 10^{-8}$ ,  $7.650410496 \cdot 10^{-10}$ ,  
 $3.187671040 \cdot 10^{-11}$ ,  $8.3886080 \cdot 10^{-13}$ ,  $1.0485760 \cdot 10^{-14}$ ]

> *plot([seq([k, 0], [k, f(k)], k = 0..20)], thickness = 2, color = red, title  
 = Binomiális eloszlás)*



$$\begin{aligned}
 &> \sum_{k=5}^{20} f(k) = \sum_{k=5}^{20} f(k) \\
 &\sum_{k=5}^{20} \left( \begin{cases} 0 & k < 0 \\ \text{binomial}(20, k) \cdot 0.2^k \cdot 0.8^{20-k} & \text{otherwise} \end{cases} \right) = 0.3703517361 \quad (1.5.10)
 \end{aligned}$$

Figyelem Maple-ben diszkrét esetben a  $CDF(X, k) = P(X \leq k)$  képlettel értelmezett, tehát az összegzésben résztvesz az utolsó k érték is!

$$\begin{aligned}
 &> 1 - CDF(X, 4) \\
 &0.3703517361 \quad (1.5.11)
 \end{aligned}$$

Tehát  $P(5 \leq X) = 1 - CDF(X, 4) = 0.37$ , vagyis 37% annak az esélye, hogy 20 izzóból 5 vagy 5-nél több izzó működik 3 hónapnál tovább.

(b) Az  $E(X) = n \cdot p = 20 \cdot 0.2 = 4$  várható érték azt jelöli, hogy átlagosan 3 hónapnál tovább, hány darab izzó működik! Ezt a 4 izzót várhatóan nem kell kicserélni, viszont a  $20 - 4 = 16$  izzó cserére szorul majd 3 hónapon belül.

$$\begin{aligned}
 &> \text{Átlagosan a működő izzók száma 3 hónap után} = \text{Mean}(X) \\
 &\text{Átlagosan a működő izzók száma 3 hónap után} = 4.0000000000 \quad (1.5.12)
 \end{aligned}$$

### 9.5.11. PÉLDA. Független ugyanolyan paraméterű Bernoulli-változók összege binomiális eloszlású

Mutassuk meg, hogy ha  $X_1, X_2, \dots, X_n$  n darab független Bernoulli-típusú valószínűségi változó, ugyanazzal a  $p$  paraméterrel, akkor ezek összege  $X = X_1 + X_2 + \dots + X_n$  binomiális eloszlású valószínűségi változó  $(n, p)$  paraméterrel.

#### MEGOLDÁS

Nézzük előbb az  $n = 2$  esetet! Legyen  $X_1$  és  $X_2$  eloszlása Bernoulli, azonos  $p$  paraméterrel

$$\begin{bmatrix} X_1 & 1 & 0 \\ P & p & 1-p \end{bmatrix} \text{ és } \begin{bmatrix} X_2 & 1 & 0 \\ P & p & 1-p \end{bmatrix}$$

Az  $X = X_1 + X_2$  összeg lehetséges értékei 0, 1 vagy 2.

Az  $X = X_1 + X_2 = 0$  csak az  $(X_1 = 0, X_2 = 0)$  esetben fordulhat elő. A függetlenség miatt  $P(X_1 = 0, X_2 = 0) = P(X_1 = 0) P(X_2 = 0) = (1-p)^2$ .

Az  $X = X_1 + X_2 = 1$  két esetben lehet, ha vagy  $(X_1 = 1, X_2 = 0)$  vagy  $(X_1 = 0, X_2 = 1)$ . A függetlenség miatt

$$P(\{X_1 = 1, X_2 = 0\} + \{X_1 = 0, X_2 = 1\}) = P(\{X_1 = 1, X_2 = 0\}) + P(\{X_1 = 0, X_2 = 1\}) = p(1-p) + (1-p)p = 2 \cdot p \cdot (1-p)$$

Az  $X = X_1 + X_2 = 2$  az  $(X_1 = 1, X_2 = 1)$  esetben fordulhat elő. A függetlenség miatt  $P(X_1 = 1, X_2 = 1) = P(X_1 = 1) P(X_2 = 1) = p^2$ .

Tehát  $X$  eloszlása

$$\begin{bmatrix} X & 0 & 1 & 2 \\ P & (1-p)^2 & 2p(1-p) & p^2 \end{bmatrix}$$

Ez az eloszlás  $n=2$  és  $p$  paraméterű binomiális valószínűségi változó eloszlása.

A feladatot  $(n-1)$ -ről  $n$ -re való teljes indukcióval bizonyíthatjuk. ♣

### 9.5.12. PÉLDA. Teszt feladatok megoldása találgatással

Egy teszt 25 kérdést tartalmaz, amelyek mindegyikére 4 lehetőség közül tudunk választani. A négy lehetőség közül pontosan egy a helyes. Tegyük fel, hogy egy diák minden kérdés esetén csak találgat.

Jelölje  $X$  a 25 kérdés közül az eltalált helyes válaszok számát!

(a) Rajzoljuk fel  $X$  eloszlását!

(b) Mekkora a valószínűsége, hogy a diák 20-nál több kérdésre jól válaszol?

(c) Mekkora a valószínűsége, hogy a diák 5-nél kevesebb kérdésre válaszol jól?

(d) Rajzoljuk fel  $X$  eloszlásfüggvényét!

(e) Várhatóan hány kérdésre válaszol jól a diák?

#### Megoldás

Legyen  $A = \{\text{egy kérdésre találgatással a diák jó választ ad}\}$  esemény, amelynek a valószínűsége  $p = P(A) = \frac{1}{4}$ , mert a diák a 4 lehetőség közül véletlenszerűen találgat!

Az  $A$  eseményre  $n=25$  független kísérletet végzünk.

Az

$X = \text{a 25 kérdésből a jó válaszok száma}$   
valószínűségi változó binomiális eloszlású lesz, mert a fenti eljárás a Bernoulli-kísérlet egy megvalósítása. Tehát  $X \sim \text{Binomial}\left(25, \frac{1}{4}\right)$

```

> restart
> with(Statistics) :
> n, p := 25, 1/4;
> X := RandomVariable(Binomial(n, p)) :

```

$$n, p := 25, \frac{1}{4}$$

(1.5.13)

Az X diszkrét változó lehetséges értékei 0,1,2,...,25. Az  $X=k$  valószínűségét a  $ProbabilityFunction(X,k)$  adja meg

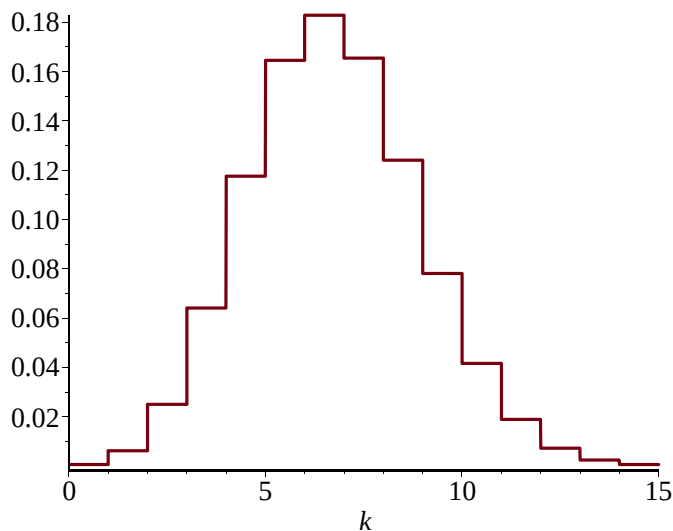
$$P(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

```

> fX := unapply(ProbabilityFunction(X, k), k) :
> fX(k);
> plot(fX(floor(k)), k = 0..15)

```

$$\begin{cases} 0 & k < 0 \\ \text{binomial}(25, k) \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{25-k} & \text{otherwise} \end{cases}$$



(b) Mekkora a valószínűsége, hogy a diák 20-nál több kérdésre jól válaszol?  
 $P(X > 20) = fX(21) + fX(22) + \dots + fX(25) = 1 - CDF(X, 20)$

```

> $\left(\sum_{k=21}^{25} fX(k) \right); \text{evalf}(\%)$

```

$$\frac{544763}{562949953421312} = 9.6769348090 \cdot 10^{-10}$$

(1.5.14)

```

> 1 - CDF(X, 20)

```

$$\frac{544763}{562949953421312}$$

(1.5.15)

Gyakorlatilag lehetetlen 20-nál több jó választ adni találgatással!

(c) Mekkora a valószínűsége, hogy a diák 5-nél kevesebb kérdésre válaszol jól?

$$P(X < 5) = fX(0) + fX(1) + \dots + fX(4) = CDF(X, 4)$$

>  $fX(0) + fX(1) + fX(2) + fX(3) + fX(4)$ ; evalf (%)

$$\frac{120325442894109}{562949953421312}$$

$$0.2137409234$$

(1.5.16)

Tehát találgatással 5-nél kevesebb jó választ kb. 21.4% eséllyel lehet elérni. Ugyanezt kapjuk a beépített CDF eloszlás függvénnyel is!

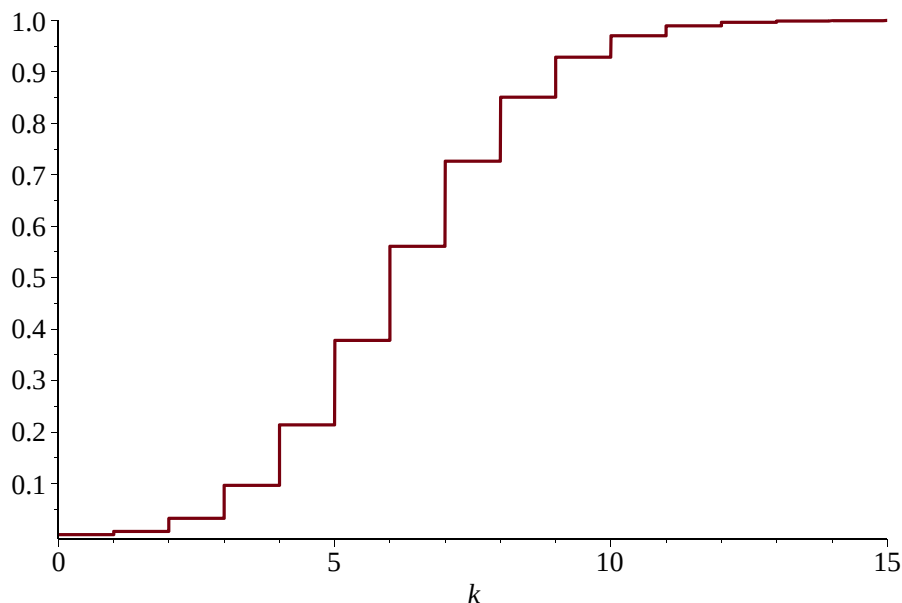
>  $CDF(X, 4)$

$$\frac{120325442894109}{562949953421312}$$

(1.5.17)

(d) Rajzoljuk fel X eloszlásfüggvényét!

>  $plot(CDF(X, floor(k)), k = 0..15)$



(e) Várhatóan hány kérdésre válaszol jól a diák?

A binomiális eloszlás várható értéke =  $n \cdot p$

> várható érték =  $n \cdot p$ ; evalf (%)

$$\text{várható érték} = \frac{25}{4}$$

$$\text{várható érték} = 6.250000000$$

(1.5.18)

>  $Mean(X)$

21.5.18

Átlagosan a jó válaszok száma találgatással 6 illetve 7 körül lesz!

## ▼ 9.6. Hipergeometrikus eloszlás

### 9.6.1. MINTAVÉTELEZÉS VISSZATEVÉS NÉLKÜL ÉS A DEFINÍCIÓ

A visszatevés nélküli mintavételezés egy másik példa a diszkrét valószínűségi eloszlásra.

Tekintsünk egy dobozt, melyben  $N$  darab alkatrész pl. csavar van.

A dobozban levő alkatrészek között  $S$  darab kitüntetett alkatrész (pl. selejtes) található ( $0 < S \leq N$ ).

Visszatevés nélkül véletlenszerűen kiválasztunk egy  $n$  elemű mintát (tegyük fel, hogy  $n \leq S$ ).

Ekkor legyen az  $X$  valószínűségi változó a kivett mintában levő selejtek száma, melynek értéke lehet:  $0, 1, 2, \dots, \min(n, S) = n$

$$P(\text{az } n \text{ kihúzott csavar között } k \text{ darab selejtes van}) = P(X=k) = \frac{\binom{S}{k} \binom{N-S}{n-k}}{\binom{N}{n}}.$$

Az  $X$  eloszlását felírhatjuk a következő táblázatos formában

|     |                                                    |                                                      |     |                                                      |     |                                                    |
|-----|----------------------------------------------------|------------------------------------------------------|-----|------------------------------------------------------|-----|----------------------------------------------------|
| $X$ | 0                                                  | 1                                                    | ... | $k$                                                  | ... | $n$                                                |
| $P$ | $\frac{\binom{S}{0} \binom{N-S}{n}}{\binom{N}{n}}$ | $\frac{\binom{S}{1} \binom{N-S}{n-1}}{\binom{N}{n}}$ | ... | $\frac{\binom{S}{k} \binom{N-S}{n-k}}{\binom{N}{n}}$ | ... | $\frac{\binom{S}{n} \binom{N-S}{0}}{\binom{N}{n}}$ |

### 9.6.2. DEFINÍCIÓ. Hipergeometrikus eloszlás

Az  $X$  valószínűségi változó **Hipergeometrikus** eloszlású  $N, n, S$  paraméterekkel, ha  $X$  lehetséges értéke  $0, 1, \dots, \min(n, S)$  egész lehet és

$$P(X=k) = \frac{\binom{S}{k} \binom{N-S}{n-k}}{\binom{N}{n}} \quad (k=0, 1, \dots, \min(n, S))$$

Ebben az esetben az  $X \sim \text{hipergeometrikus}(N, S, n)$  jelölést használjuk.

Képezzük a fenti formulát szimbolikus módon, adott  $N, S, n$  és  $k$  esetén!

> *restart; with(plots) : with(combinat) : with(Statistics) :*

> *hipergeometric := (N, S, n, k) →  $\frac{\text{binomial}(S, k) \text{binomial}(N-S, n-k)}{\text{binomial}(N, n)}$*

*hipergeometric := (N, S, n, k)*

(1.6.1)

*→  $\frac{\text{combinat:binomial}(S, k) \text{combinat:binomial}(N-S, n-k)}{\text{combinat:binomial}(N, n)}$*

Ugyanezt a képletet megadja a "Statistics" csomagban található "Hypergeometric(N,S,n)" eljárás is!

> *X := RandomVariable(Hypergeometric(N, S, n)) :*

> *ProbabilityFunction(X, k)*

$$\left\{ \begin{array}{ll} 0 & k < 0 \\ 0 & S < k \\ \frac{\text{binomial}(S, k) \text{binomial}(N - S, n - k)}{\text{binomial}(N, n)} & \text{otherwise} \end{array} \right. \quad (1.6.2)$$

Adjuk meg a  $0 < n \leq S < N$  korlátozó feltételeket és adjuk össze a valószínűségeket!

> *assume(S < N); additionally(n ≤ S); additionally(0 < S) : additionally(0 < n) :*

$$\sum_{k=0}^n \text{hipergeometric}(N, S, n, k) = \sum_{k=0}^n \text{hipergeometric}(N, S, n, k)$$

$$\sum_{k=0}^{n\sim} \frac{\text{binomial}(S\sim, k) \text{binomial}(N\sim - S\sim, n\sim - k)}{\text{binomial}(N\sim, n\sim)} = 1 \quad (1.6.3)$$

Ezzel a formulával valóban diszkrét eloszlást adtunk meg, mert összegezve a valószínűségeket  $k = 0$  tól  $n$ -ig 1-t kapunk

$$\sum_{k=0}^n \frac{\binom{S}{k} \binom{N-S}{n-k}}{\binom{N}{n}} = 1, \text{ ha } n \leq S.$$

### 9.6.3. TÉTEL. VÁRHATÓ ÉRTÉK, VARIANCIA

Az  $X \sim \text{hipergeometrikus}(N, S, n)$  hipergeometrikus eloszlás **várható értéke**

$$\mu = n \cdot \frac{S}{N}$$

**varianciája**

$$\sigma^2 = n \cdot \frac{S}{N} \cdot \left(1 - \frac{S}{N}\right) \left(1 - \frac{n-1}{N-1}\right)$$

### Bizonyítás

A Maple beépített algebrai egyszerűsítési képességeit használva mutatjuk meg a várható érték és variancia számítását.

> *Sum(k·hipergeometric(N, S, n, k), k = 0..n); convert(%, factorial);*  
*Várható\_érték := simplify(value(%));*

$$\sum_{k=0}^{n\sim} \frac{k \text{binomial}(S\sim, k) \text{binomial}(N\sim - S\sim, n\sim - k)}{\text{binomial}(N\sim, n\sim)}$$

$$\sum_{k=0}^{n\sim} \frac{k S\sim! (N\sim - S\sim)! n\sim! (N\sim - n\sim)!}{k! (S\sim - k)! (n\sim - k)! (N\sim - S\sim - n\sim + k)! N\sim!}$$

$$\text{Várható_érték} := \frac{S\sim n\sim}{N\sim} \quad (1.6.4)$$

> *Mean(X);*

$$\frac{S\sim n\sim}{N\sim} \quad (1.6.5)$$

> *Sum(k^2·hipergeometric(N, S, n, k), k = 0..n) : % = simplify(value(convert(%, factorial))) ;*

$Sz\acute{o}r\acute{a}s\_n\acute{e}gyszet = factor(rhs(\%) - V\acute{a}rható\_\acute{e}rt\acute{e}k^2);$

$$\sum_{k=0}^{n\sim} \frac{k^2 \text{binomial}(S\sim, k) \text{binomial}(N\sim - S\sim, n\sim - k)}{\text{binomial}(N\sim, n\sim)} = \frac{S\sim n\sim (S\sim n\sim + N\sim - S\sim - n\sim)}{(N\sim - 1) N\sim}$$

$$Sz\acute{o}r\acute{a}s\_n\acute{e}gyszet = \frac{n\sim S\sim (N\sim - S\sim) (N\sim - n\sim)}{(N\sim - 1) N\sim^2} \quad (1.6.6)$$

> *Variance(X)*

$$\frac{n\sim S\sim \left(1 - \frac{S\sim}{N\sim}\right) (N\sim - n\sim)}{N\sim (N\sim - 1)} \quad (1.6.7)$$

Q.e.d.

#### 9.6.4. A HISZTOGRAM VAGY ELOSZLÁS DIAGRAMJA

A hipergeometrikus eloszlás hisztogramjának ábrázolására készítünk egy eljárását, amelynek paraméterei N, S és n.

> *with(plots) :*

*HipHist := proc(N, S, n)*

*local f, k, c, d, h, P, Q;*

*c := blue; P := NULL : Q := NULL :*

*for k from 0 to n do*

*h := binomial(S, k) · binomial(N - S, n - k) / binomial(N, n);*

*P := P, plot( [[k-.5, 0], [k+.5, 0], [k+.5, h], [k-.5, h], [k-.5, 0]], color = c );*

*if (k < n/2)*

*then Q := Q, plot( [[k+.5, 0], [k+.5, h]],*  
*color = c) :*

*else Q := Q, plot( [[k-.5, 0], [k-.5, h]],*  
*color = c) : fi;*

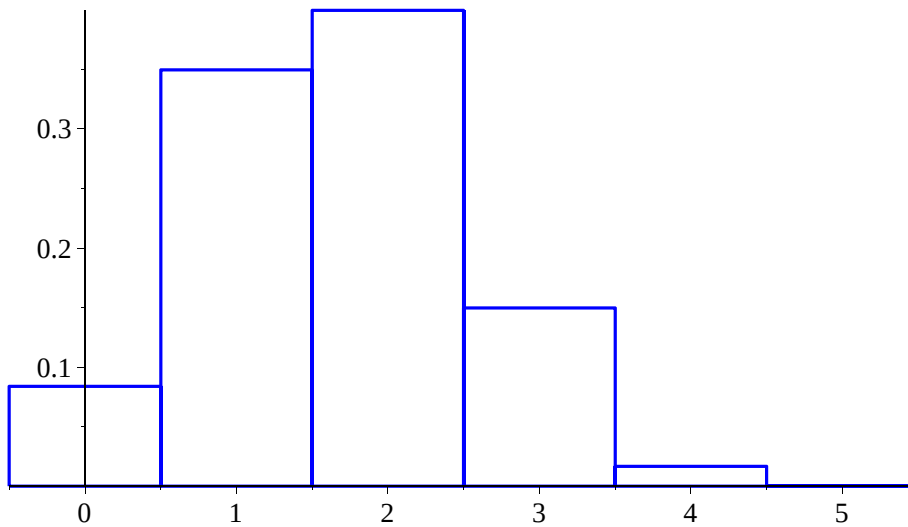
*od;*

*plots[display]( [ P, Q ] );*

*end proc;*

A hisztogramot nézzük meg N = 15, S = 5 és n = 5 esetére!

> *HipHist( 15, 5, 5)*



#### 9.6.5. HISZTOGRAM ANIMÁCIÓK VÁLTOZÓ PARAMÉTEREK MELLETT

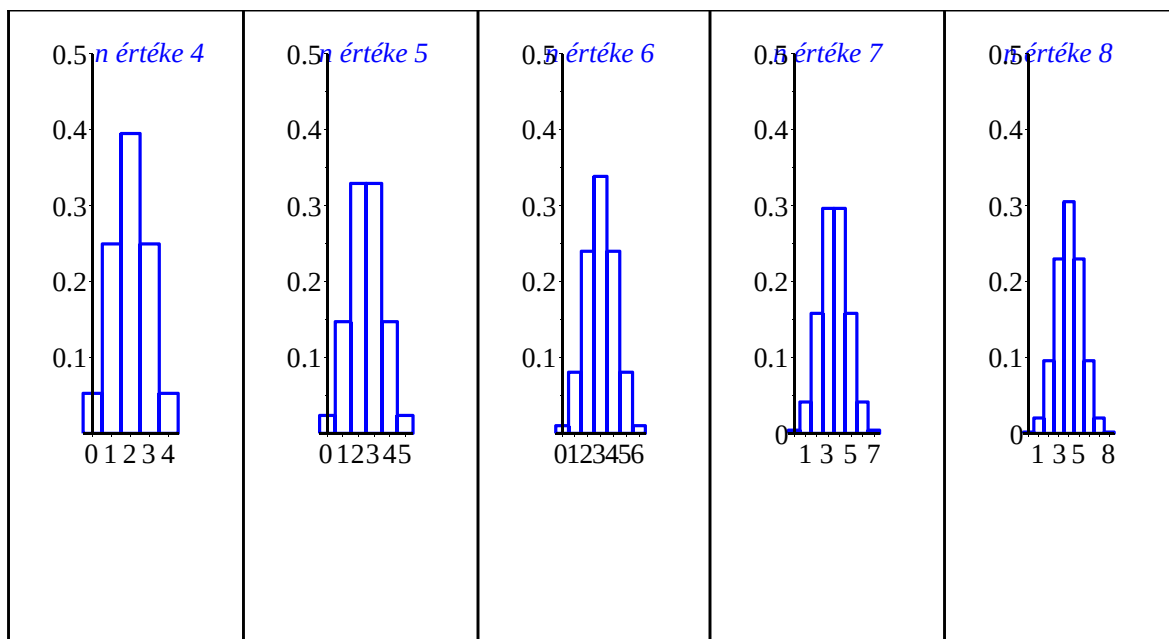
Vizsgáljuk az eloszlás alakját és formáját állandó  $N = 40$  és  $S=20$  esetén és növekvő  $n$  esetén  $n = 4$ -től  $n = 8$ -ig!

```
> N := 40; S := 20; P := NULL :
> for n from 4 to 8 do
> Hn := HipHist(N, S, n) :
> num := convert(n, string) :
> trackern := textplot([3, 0.5, cat(n értéke , num)], color = blue) :
> P := P, display({ Hn, trackern }) :
> od :
> display(Matrix(1, 5, [P]))
```

$N := 40$

$S := 20$





A következő alkalmazás számításai és szemléltetése számítógép algebrai rendszer nélkül szegényes lenne.

#### 9.6.6. PÉLDA. Tóban élő halak száma

Egy halgazdaságban úgy becsülik meg a halak számát, hogy kifognak bizonyos mennyiségű halat, megjelölik és visszadobják a tóba. Néhány nap múlva, amikor a megjelölt halak elkeveredtek a többiek között, akkor újból kifognak néhány halat és megszámlálják, hogy hány jelölt van közöttük.

Tegyük fel, hogy  $m=50$  halat fogtak ki először és ezek mindegyikét megjelölték. A később kifogott halak száma  $n=120$  volt, amelyek között  $k=10$  megjelöltet találtak.

(a) Mennyi a valószínűsége, hogy  $N=800$  hal él a tóban?

(b) Rajzoljuk fel az  $(N, m, n)$  paraméterű hipergeometrikus eloszlásnak a  $k=10$  értékhez tartozó valószínűségeit különböző  $N$  halszám függvényében?

(c) A halászok azt mondják, hogy a halak száma olyan  $N$  érték, melyre a (b) pontban kapott valószínűség a maximális! Határozzuk meg ilyen módon a halak  $N$  számát a tóban és ennek mekkora a valószínűsége?

#### Megoldás

(a) Felvesszük az adatokat

> restart

>  $m, n, N := 50, 120, 800$

$m, n, N := 50, 120, 800$

(1.6.8)

Ha az  $X$  valószínűségi változó jelöli a másodszorra kifogott 120 hal között a megjelölt halak számát, akkor  $X$  eloszlása hipergeometrikus és

$$P(X = k) = \frac{\binom{m}{k} \cdot \binom{N-m}{n-k}}{\binom{N}{n}}$$

$$X \sim \text{Hypergeometric}(N, m, n)$$

Mert a sokaság száma  $N = 800$ , amelyből a halakat kivesszük. A megjelöltek száma (megfelelnek a selejteknék)  $m = 50$  és  $n = 120$  elemű mintát veszünk ki. Keressük az eloszlás  $P(k = 10)$  valószínűsége!

```

> with(Statistics) :
> X := RandomVariable(Hypergeometric(N, m, n)) :
> fX := unapply(ProbabilityFunction(X, k), k) :
> p = evalf(fX(10))

```

$$p = 0.0897579305 \quad (1.6.9)$$

Tehát kb. 8.9% annak esélye, hogy a tóban N=800 hal van!

(b) Az N értékét Maple-ben felszabítjuk és felrajzoljuk az (N, m, n) paraméterű hipergeometrikus eloszlásnak a k=10 értékhez tartozó valószínűségeit az N halszám függvényében? a

```

> N := 'N':
> Y := RandomVariable(Hypergeometric(N, m, n))
> fY := unapply(ProbabilityFunction(Y, k), k)

```

$Y := \_R7$

$$fY := k \rightarrow \text{piecewise}\left(k < 0, 0, 50 < k, 0, \frac{\text{binomial}(50, k) \text{binomial}(N - 50, 120 - k)}{\text{binomial}(N, 120)}\right) \quad (1.6.10)$$

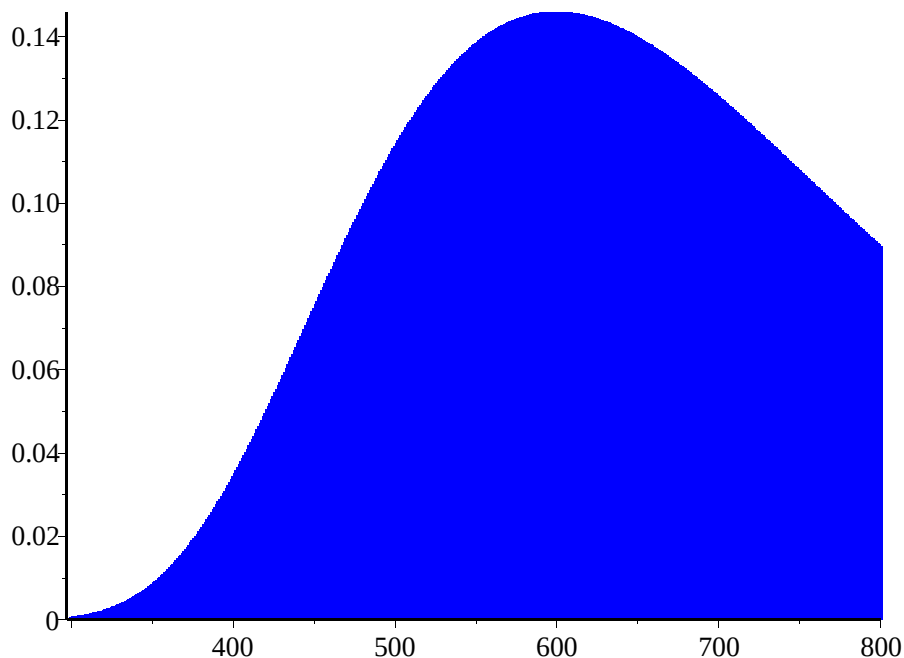
```
> fY(10)
```

$$\frac{10272278170 \text{binomial}(N - 50, 110)}{\text{binomial}(N, 120)} \quad (1.6.11)$$

```

> palcikak := [seq([N, 0], [N, evalf(fY(10))], N = 300..800)] :
> plot(palcikak, color = blue)

```



Elénk tárult egy olyan változás, amely a tóban élő halak N számának növekedésével mutatja annak valószínűségét, hogy a kifogott n=120 hal között k=10 megjelölt van. Kezdetben N

növekedésével nő a valószínűség, majd eléri maximumát  $N=600$ -nál és onnan kezdve csökken. Tehát legnagyobb valószínűséggel 600 hal van a tóban!

> legnagyobb valószínűség = evalf( eval( fY( 10 ), N = 600 ) )

$$\text{legnagyobb valószínűség} = 0.1460265080 \quad (1.6.12)$$

Ez a maximális valószínűség 14,6%.

## 9.7. Poisson eloszlás

Poisson eloszlást használunk, ha a **binomiális eloszlás**ban  $n$  értéke "nagy" és ugyanakkor  $p$  értéke "kicsi". Ilyen esetekben a binomiális eloszlás értékeinek számítása idő igényes és pontatlan, míg a Poisson közelítés  $\lambda = n p$  paraméter értékkel elfogadható hibával közelít és értékei gyorsan számolhatók.

Poisson eloszlással modellezhető az ügyfelek érkezése bankba, boltba stb., általában egy **várakozási sorba**.

A Poisson-eloszlás alapul szolgál az **exponenciális eloszláshoz**. Ugyanis, ha adott időtartam alatt érkező ügyfelek száma Poisson eloszlást követ, akkor az érkezők között eltelt időtartam exponenciális eloszlású lesz.

### 9.7.1. DEFINÍCIÓ. Poisson-eloszlás

Az  $X$  valószínűségi változót, melynek értéke 0, 1, 2, ... lehet és az egyes értékekhez tartozó valószínűségeket az alábbiak szerint számítjuk

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, \dots \text{ és } 0 < \lambda,$$

akkor  $X$ -t  $\lambda$ -paraméterű Poisson-eloszlású diszkrét valószínűségi változónak nevezzük.

Jelölése:  $X \sim \text{Poisson}(\lambda)$ .

Képezzük a valószínűségeket szimbolikus módon, adott  $k$  és  $0 < \lambda$  esetén!

$$> \text{restart : Poison} := k \rightarrow \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\text{Poison} := k \rightarrow \frac{\lambda^k e^{-\lambda}}{k!} \quad (1.7.1)$$

Ezzel a formulával valóban diszkrét eloszlást adtunk meg, mert összegezve a valószínűségeket 1-t kapunk.

$$> \sum_{k=0}^{\infty} \text{Poison}(k) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = 1 \quad (1.7.2)$$

Az összegzés alapja a

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots + \frac{\lambda^n}{n!} + \dots = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \quad (\text{tetszőleges valós } \lambda \text{ esetén})$$

konvergens Taylor-sor. Ez alapján az alábbi végtelen összeg valóban 1 lesz

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \left( \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) = e^{-\lambda} e^{\lambda} = 1$$

A *Statistics* csomagban a  $\text{Poisson}(\lambda)$  nevű eljárással tudunk Poisson-eloszlású változót létrehozni.

>  $\text{assume}(0 < \lambda) : \text{with}(\text{Statistics}) : X := \text{RandomVariable}(\text{Poisson}(\lambda)) :$   
 >  $f := \text{unapply}(\text{ProbabilityFunction}(X, k), k) :$   
 $P('X' = k) = f(k)$

$$P(X = k) = \begin{cases} 0 & k < 0 \\ \frac{\lambda^k e^{-\lambda}}{k!} & \text{otherwise} \end{cases} \quad (1.7.3)$$

### 9.7.2. TÉTEL. Poisson-eloszlás származtatása binomiálisból

Legyen  $X_n \sim \text{binomiális}(n, p_n)$ , ahol  $n \geq 1$  egész. Tegyük fel, hogy  $n \cdot p_n = \lambda$ , ahol  $\lambda > 0$ . Legyen  $Y \sim \text{Poisson}(\lambda)$  Poisson-eloszlású változó. Ekkor

$$\lim_{n \rightarrow \infty} P(X_n = k) = P(Y = k)$$

minden  $k = 0, 1, 2, \dots$  egész esetén.

#### Bizonyítás

Tekintsük az egyszerű  $k = 0$  esetet, mikor az

$$P(X_n = 0) = (1 - p_n)^n = \left(1 - \frac{n \cdot p_n}{n}\right)^n = \left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda} = P(Y = 0),$$

ahol  $n$  tart végtelenhez.

Tekintsük a  $k = 1$  esetet. Ekkor

$$P(X_n = 1) = n \cdot p_n \cdot (1 - p_n)^{n-1} = \lambda \cdot \frac{(1 - p_n)^n}{1 - p_n} \rightarrow \lambda \cdot \frac{e^{-\lambda}}{1} = P(Y = 1),$$

ha  $n$  tart végtelenhez, mert  $p_n = \frac{\lambda}{n} \rightarrow 0$ .

Ugyanígyen átalakításokkal kapjuk tetszőleges  $k$  egész mellett

$$\begin{aligned} P(X_n = k) &= \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} = \frac{1}{k!} n \cdot (n-1) \cdots (n-k+1) \cdot p_n^k (1 - p_n)^n \cdot \frac{1}{(1 - p_n)^k} = \\ &= \frac{(n \cdot p_n)^k}{k!} \cdot 1 \cdot \frac{(n-1)}{n} \cdots \frac{(n-k+1)}{n} \cdot (1 - p_n)^n \cdot \frac{1}{(1 - p_n)^k} \rightarrow \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda}, \end{aligned}$$

ha  $n$  tart végtelenbe. Ahol felhasználtuk az  $n \cdot p_n = \lambda$  egyenlőséget, az

$$\frac{(n-i)}{n} \rightarrow 1 \quad (n \rightarrow \infty, \text{fix } i - re) \text{ határértéket és a korábban igazolt } (1 - p_n)^n \rightarrow e^{-\lambda} \quad (n \rightarrow \infty)$$

határérték relációt. Q.e.d.

A  $n \cdot p_n = \lambda$  egyenlőség helyettesíthető a gyengébb  $n \cdot p_n \rightarrow \lambda \quad (n \rightarrow \infty)$  feltétellel.

### 9.7.3. TÉTEL. Poisson-eloszlás várható értéke és varianciája

Ha  $X \sim \text{Poisson}(\lambda)$ , akkor a várható értéke és varianciája is a paraméter

$$E(X) = \lambda, \text{Var}(X) = \lambda$$

#### Bizonyítás

Megmutatjuk, hogy a  $\lambda$  paraméterű Poisson- eloszlás **várható értéke**  $E(X) = \lambda$ .

A várható érték számítási szabálya alapján

$$E(X) = \sum_{k=0}^{\infty} x_k \cdot P(X=x_k) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \stackrel{k-1=i}{=} e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^{i+1}}{i!} = \lambda.$$

Ugyanezt a Maple szimbolika is magadja.

$$> \sum_{k=0}^{\infty} k f(k) = \sum_{k=0}^{\infty} k f(k); \text{Várható\_érték} := \text{Mean}(X)$$

$$\sum_{k=0}^{\infty} k \left( \begin{cases} 0 & k < 0 \\ \frac{\lambda^k e^{-\lambda}}{k!} & \text{otherwise} \end{cases} \right) = \lambda$$

$$\text{Várható\_érték} := \lambda$$

(1.7.4)

Megmutatjuk, hogy a Poisson- eloszlás **szórásnégyzete** =  $\sigma^2 = \lambda$

$$> \sum_{k=0}^{\infty} k^2 f(k) = \sum_{k=0}^{\infty} k^2 f(k); \text{Szórás\_négyzet} = \text{factor}(\text{simplify}(\text{rhs}(\%)) - \text{Várható\_érték}^2)$$

$$\sum_{k=0}^{\infty} k^2 \left( \begin{cases} 0 & k < 0 \\ \frac{\lambda^k e^{-\lambda}}{k!} & \text{otherwise} \end{cases} \right) = \lambda (\lambda + 1)$$

$$\text{Szórás\_négyzet} = \lambda$$

(1.7.5)

$$> \text{Variance} := \text{Variance}(X)$$

$$\text{Variance} := \lambda$$

(1.7.6)

Q.e.d.

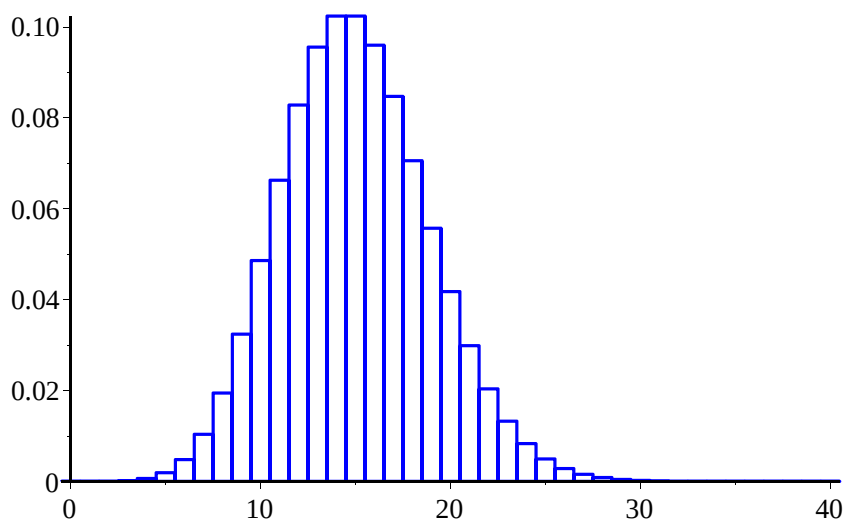
#### 9.7.4. HISZTOGRAM, ELOSZLÁS RAJZOLÁS

Poisson eloszlás hisztogramjának vagy eloszlásának ábrázolására készítünk egy eljárását, amelynek paramétere a  $\lambda$  és  $n$ .

```
> PoissonHist := proc(lambda, n)
 local f, k, c, d, h, P, Q;
 c := blue;
 P := NULL;
 Q := NULL;
 for k from 0 to n do
 h := lambda^k * exp(- lambda) / factorial(k);
 P := P, plot([[k - Float(5, -1), 0], [k + Float(5, -1), 0], [k + Float(5, -1), h], [k - Float(5, -1), h], [k - Float(5, -1), 0]], color = c);
 if k < n/2 then
 Q := Q, plot([[k + Float(5, -1), 0], [k + Float(5, -1), h]], color = c)
 else
 Q := Q, plot([[k - Float(5, -1), 0], [k - Float(5, -1), h]], color = c)
 end if
 end do;
 plots[display]([P, Q])
end proc :
```

A hisztogramot nézzük  $\lambda = 2$  és  $n = 40$ -re!

> *PoissonHist*( 15, 40)



#### 9.7.5. HISZTOGRAM ANIMÁCIÓJA VÁLTOZÓ $\lambda$ PARAMÉTER MELLETT

Vizsgáljuk az eloszlás alakját és formáját növekvő  $\lambda$  esetén  $\lambda = 3$  -tól  $\lambda = 8$ -ig!

$n := 30$  : *with*(*plots*) :  $P := \text{NULL}$  :

> **for** **lambda** **from** 3 **to** 8 **do**

>  $H_\lambda := \text{PoissonHist}(\lambda, n)$  :

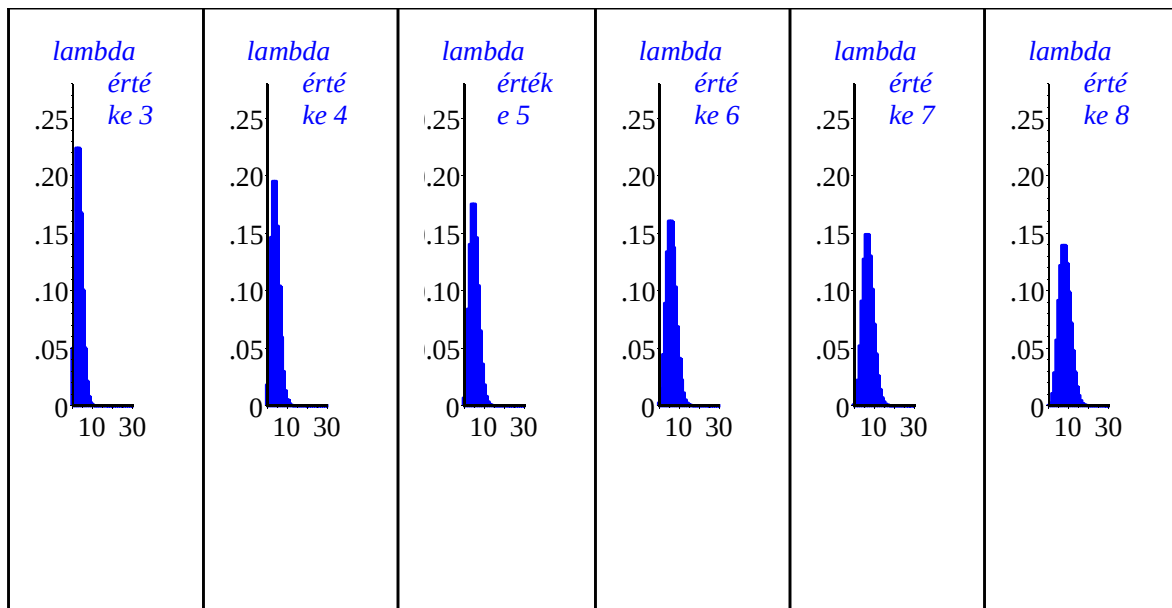
>  $\text{num} := \text{convert}(\lambda, \text{string})$  :

>  $\text{tracker}_\lambda := \text{textplot}\left(\left[\frac{n}{2}, 0.28, \text{cat}(\text{lambda értéke}, \text{num})\right], \text{color} = \text{blue}\right)$  :

>  $P := P, \text{display}(\{H_\lambda, \text{tracker}_\lambda\})$  :

> **od** :

>  $\text{display}(\text{Matrix}(1, 6, [P]))$



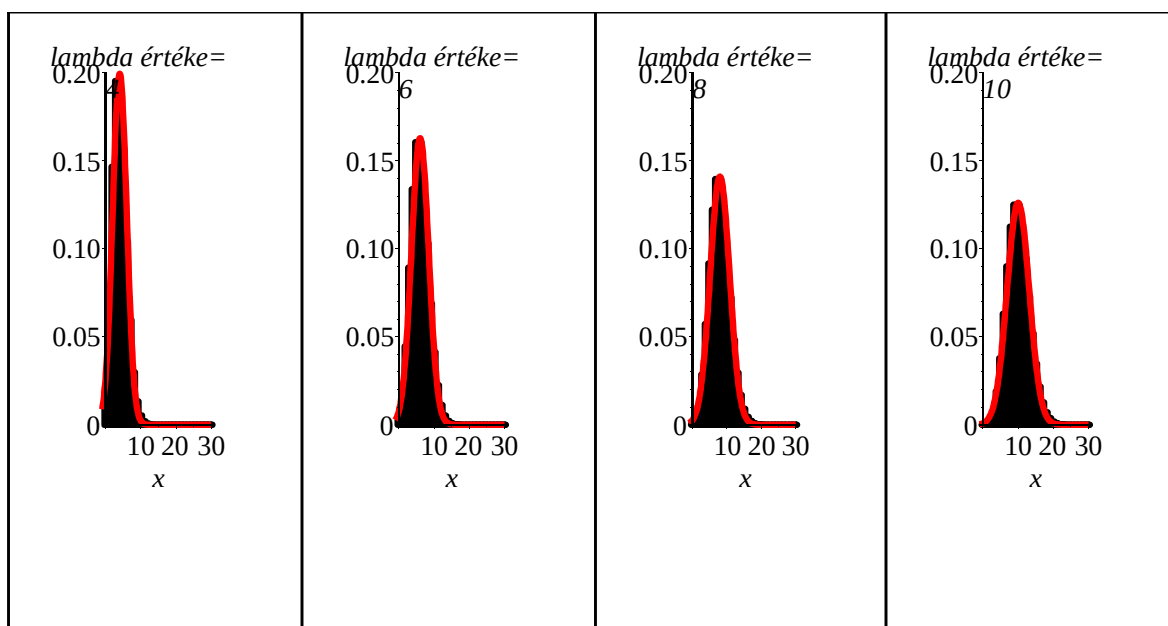
Látható, hogy  $\lambda$  növekedésével az eloszlás maximum helye jobbra vándorol, mert a várható érték a paraméter. Továbbá az eloszlás szórása is nagyobb lesz, mert a variancia is a paraméter. Az eloszlás alakja egyre jobban hasonlít egy normál eloszlásra. Ebben a centrális határeloszlás tétel eredménye jelenik meg.

#### 9.7.6. POISSON ELOSZLÁS KÖZELÍTÉSE NORMÁL ELOSZLÁSSAL

Szemléltessük a fenti nevezetes eredményt, mely szerint a  $\lambda$  paraméter növelésével a Poisson( $\lambda$ ) eloszlás az  $N(\lambda, \sqrt{\lambda})$  normál eloszlás sűrűség függvényét egyre jobban megközelíti.

Fix  $\lambda = 4$ -től 10-ig kettesével növekedve felrajzoljuk a Poisson( $\lambda$ ) binomiális eloszlás hisztogramját és az ugyanolyan várható értékű és szórású normál eloszlást!

```
> restart : with(Statistics) : with(plots) :
> kepek := NULL :
 for λ from 4 to 10 by 2 do
> kep1 := ColumnGraph([seq(ProbabilityFunction(Poisson(λ), k), k = 0 ..30)], color
 = blue, thickness = 3, width = 0.7, distance = 0.3, offset = -0.35) :
> kep2 := plot(PDF(Normal(λ , $\sqrt{\lambda}$), x), x = -1 ..30, color = red, thickness = 3) :
> num := convert(λ , string) : szöveg := textplot([10, 0.2, cat(lambda értéke = , num)], color
 = black) :
> kepek := kepek, display([kep1, kep2, szöveg]) :
 end do:
> display(Matrix(1, 4, [kepek]))
```



Valóban  $\lambda$  növekedésével a  $\text{Poisson}(\lambda)$  eloszlás egyre jobban megközelíti a  $\text{Normal}(\mu = \lambda, \sigma = \sqrt{\lambda})$  normál eloszlást ugyanolyan várható értékkel és szórással, mint a Poisson- eloszlás. A Poisson-eloszlás alkalmazhatósága nagyrészt múlik azon is, hogy független Poisson-eloszlások összege is Poisson-eloszlás lesz.

#### 9.7.7. FÜGGETLEN POISSON ELOSZLÁSOK ÖSSZEGE POISSON ELOSZLÁS

Igazoljuk, hogy ha  $X_1 \sim \text{Poisson}(\lambda_1)$  és  $X_2 \sim \text{Poisson}(\lambda_2)$  függetlenek, akkor az  $Y = X_1 + X_2$  valószínűségi változóra  $Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$ .

##### Bizonyítás

Próbáljuk előbb "kézzel" igazolni, majd a Statistics csomag segítségével is!

Mivel  $X_1$  és  $X_2$  értékei csak 0, 1, 2,... egészek lehetnek, ezért az  $Y = X_1 + X_2$  összeg lehetséges értékei is 0, 1, 2,... lehet.

Számoljuk ki a  $P(Y = n) = P(X_1 + X_2 = n)$  valószínűségeket tetszőleges  $n = 0, 1, 2, \dots$  esetén!

Az  $A = \{(X_1; X_2) | X_1 + X_2 = n\}$  esemény az alábbi  $(X_1; X_2)$  párokból áll

$$A = \{(0; n), (1; n-1), (2; n-2), \dots, (n-1; 1), n; 0\}$$

ahol a számpárok első eleme  $X_1$  értékét jelöli és a második eleme  $X_2$  értékét. Az összes lehetséges párt felsoroltuk amelyek összege  $n$  lesz, mert  $X_1$  és  $X_2$  értékei csak 0, 1, 2,... egészek lehetnek. Az  $A$  esemény valószínűségét a benne szereplő elemi események valószínűségeinek összege adja

$$P(A) = \sum_{k=0}^n P(X_1 = k, X_2 = n - k).$$

Mivel  $X_1$  és  $X_2$  változók függetlenek, ezért az együttes események valószínűségét az tényezők valószínűségeinek szorzatával számíthatjuk! Tehát

$$\begin{aligned} P(A) &= \sum_{k=0}^n P(X_1 = k, X_2 = n - k) = \sum_{k=0}^n P(X_1 = k) \cdot P(X_2 = n - k) = \sum_{k=0}^n \frac{\lambda_1^k e^{-\lambda_1}}{k!} \cdot \frac{\lambda_2^{n-k} e^{-\lambda_2}}{(n-k)!} = \\ &= \frac{1}{n!} \cdot e^{-\lambda_1} \cdot e^{-\lambda_2} \sum_{k=0}^n \frac{n!}{k! \cdot (n-k)!} \lambda_1^k \cdot \lambda_2^{n-k} = \frac{1}{n!} \cdot e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^n, \end{aligned}$$



ahol az utolsó lépésben alkalmaztuk a  $(\lambda_1 + \lambda_2)^n$  összeg binomiális-tétel szerinti kifejtését. Ezzel igazoltuk "kézzel" a tétel állítását. Nézzük a Maple megoldást!

> restart; with(Statistics) :

> X1 := RandomVariable(Poisson( $\lambda_1$ )) : X2 := RandomVariable(Poisson( $\lambda_2$ )) :

> f1 := unapply(ProbabilityFunction(X1, n), n) :

f2 := unapply(ProbabilityFunction(X2, n), n) : f1'(n) = f1(n), f2'(n) = f2(n);

$$f_1(n) = \begin{cases} 0 & n < 0 \\ \frac{\lambda_1^n e^{-\lambda_1}}{n!} & \text{otherwise} \end{cases}, f_2(n) = \begin{cases} 0 & n < 0 \\ \frac{\lambda_2^n e^{-\lambda_2}}{n!} & \text{otherwise} \end{cases} \quad (1.7.7)$$

Az összeg eloszlását a Maple sajnos **nem adja meg!**

> Y := X1 + X2 :

> ProbabilityFunction(Y, k)

FAIL (1.7.8)

> sum(f1(k) · f2(n-k), k = 0 .. n);

$$\sum_{k=0}^n \left( \begin{cases} 0 & k < 0 \\ \frac{\lambda_1^k e^{-\lambda_1}}{k!} & \text{otherwise} \end{cases} \right) \left( \begin{cases} 0 & n-k < 0 \\ \frac{\lambda_2^{(n-k)} e^{-\lambda_2}}{(n-k)!} & \text{otherwise} \end{cases} \right) \quad (1.7.9)$$

Az összeget írjuk fel elágazás nélkül, hogy a Maple képes legyen az összeget szimbolikus formában is megtalálni!

$$> \sum_{k=0}^n \frac{\lambda_1^k e^{-\lambda_1} \lambda_2^{n-k} e^{-\lambda_2}}{k! (n-k)!} = \sum_{k=0}^n \frac{\lambda_1^k e^{-\lambda_1} \lambda_2^{n-k} e^{-\lambda_2}}{k! (n-k)!}$$

> convert(simplify(%), factorial)

$$\sum_{k=0}^n \frac{\lambda_1^k e^{-\lambda_1} \lambda_2^{(n-k)} e^{-\lambda_2}}{k! (n-k)!} = \frac{e^{-\lambda_1} e^{-\lambda_2} (\lambda_1 + \lambda_2)^n}{\Gamma(n+1)}$$

$$e^{-\lambda_1 - \lambda_2} \left( \sum_{k=0}^n \frac{\lambda_1^k \lambda_2^{(n-k)}}{k! (n-k)!} \right) = \frac{e^{-\lambda_1 - \lambda_2} (\lambda_1 + \lambda_2)^n}{n!} \quad (1.7.10)$$

A szimbolikus összegzés sikeres. Q.e.d.

### 9.7.8. PÉLDA. Selejtekt gyártása

Egy CD lemezeket gyártó üzem naponta átlagosan 4 ( $=\lambda$ ) darab selejtet gyárt.

Egy két napos periódus alatt, mekkora a valószínűsége, hogy a selejtes CD-k száma nem haladja meg a 3 darabot?

### Megoldás

Mivel az átlagos selejtszám van megadva, mely 4 és a Poisson-eloszlás paramétere éppen a

várható érték, ezért az 1 nap alatt gyártott selejtek száma Poisson(4) eloszlású.

Az időtartam, amelyre a kérdés vonatkozik nem 1 nap, hanem 2 nap. Ezért vezessük be a következő valószínűségi változókat:

$X_1$  = az első nap gyártott selejtes CD-k száma

$X_2$  = a második nap gyártott selejtes CD-k száma

A feltételek szerint  $X_1 \sim \text{Poisson}(4)$  és  $X_2 \sim \text{Poisson}(4)$ , továbbá  $X_1$  és  $X_2$  függetlenek. Ezért a 9.7.7. tétel alapján az  $Y = X_1 + X_2$  összeg valószínűségi változó  $Y \sim \text{Poisson}(8)$ , mely a két nap alatt összesen gyártott selejtes CD-k száma.

> *restart : with(Statistics) :*

> *Y := RandomVariable(Poisson(8)) :*

> *p := unapply(ProbabilityFunction(Y, k), k)*

$$p := k \rightarrow \text{piecewise}\left(k < 0, 0, \frac{8^k e^{-8}}{k!}\right) \quad (1.7.11)$$

Keressük  $P(Y \leq 3)$  valószínűségét! Nézzük ezt az eloszlással!

> *P('Y' ≤ 3) = p(0) + p(1) + p(2) + p(3);*

> *evalf(%)*

$$P(Y \leq 3) = \frac{379}{3} e^{-8}$$
$$P(_R \leq 3) = 0.0423801120 \quad (1.7.12)$$

Ugyanezt adja a  $CDF(Y, k)$  eloszlás függvény értéke a  $k=3$  helyen vagy a  $Probability(Y \leq 3)$  eljárás hívás is.

> *CDF(Y, 3);*

> *Probability(Y ≤ 3)*

$$\frac{379}{3} e^{-8}$$
$$\frac{379}{3} e^{-8} \quad (1.7.13)$$

Tehát kb. 4.24% annak az esélye, hogy 2 nap alatt nem lesz több selejt, mint 3. ♣

### 9.7.9. Poisson -számítás és a binomiális számítás sebessége

Egy digitális adatátviteli rendszerben egy bit tévesztésének valószínűsége  $p = 10^{-6} = 0.000001$ . Határozzuk meg annak valószínűségét, hogy  $10^6 = 1000000$  bitből 3 vagy több bit lesz hibás! Használjunk a közelítéshez Poisson -eloszlást!

#### Megoldás

> *restart :*

> *with(Statistics) :*

Valójában binomiális eloszlást kellene használnunk  $p = 10^{-6}$  és  $n = 10^6$  paraméterekkel.

Kipróbálhatja az olvasó, hogy a megadott valószínűséget a Maple is "nagyon" sokáig számolja! A megoldás végén szerplő megjegyzést ki kell venni ehhez. Ezért Poisson eloszlást használunk helyette, melynek várható értéke megegyezik a binomiális eloszlás várható értékével!

*p, n := 10<sup>-6</sup>, 10<sup>6</sup>*

$$p, n := \frac{1}{1000000}, 1000000 \quad (1.7.14)$$

Mivel a binomiális eloszlás várható értéke  $n \cdot p$ , és a Poisson -eloszlás várható-értéke a  $\lambda$  paramétere, ezért  $\lambda = n p$  kell, hogy teljesüljön.

>  $\lambda := n p$

$$\lambda := 1 \quad (1.7.15)$$

Ezzel a  $\lambda=1$  paraméterrel létrehozuk az X Poisson - eloszlású valószínűségi változót.

$X := \text{RandomVariable}(\text{Poisson}(\lambda)) :$

$P := \text{unapply}(\text{ProbabilityFunction}(X, u), u) : 'P(X=u)' = P(u)$

$$P(X=u) = \begin{cases} 0 & u < 0 \\ \frac{e^{-1}}{u!} & \text{otherwise} \end{cases} \quad (1.7.16)$$

A feladat a  $P(X \geq 3)$  valószínűséget kérdezi.

>  $'P(X \geq 3)' = 1 - P(0) - P(1) - P(2); 'P(X \geq 3)' = \text{evalf}(\text{rhs}(\%))$

$$P(X \geq 3) = 1 - \frac{5}{2} e^{-1}$$

$$P(X \geq 3) = 0.0803013970 \quad (1.7.17)$$

Tehát 1 000 000 bitből 3 vagy 3-nál több bit tévesztésének valószínűsége kb. 8%.

Binomiális eloszlással a következőképpen számolhatunk! A következő utasítást csak akkor indítsuk el, ha "sok" időnk van!

#  $1\text{-CDF}(\text{Binomial}(n,p),2);$

#### 9.7.10. PÉLDA. Repülőtéri forgalomirányítás

Egy forgalmas repülőtéren automatikus légi forgalom irányító berendezés felügyeli a repülőgépek le - és felszállását. Egy irányító berendezés óránként maximálisan 20 repülőgép le vagy felszállását képes felügyelni. A repülőtéren átlagosan két percenként leszáll vagy felszáll egy repülőgép. Feltesszük, hogy a le- és felszálló repülőgépek számának változása Poisson-eloszlást mutat.

(a) Rajzoljuk fel a le- és felszálló repülőgépek számának eloszlását, amely egy órára vonatkozik!

(b) Az esetek hány százalékában szükséges két irányító berendezés párhuzamos működtetése egy egy órás periódus alatt?

(c) Milyen gyakran nem elegendő még két párhuzamosan működő irányító berendezés sem?

#### MEGOLDÁS

(a) Rajzoljuk fel a le- és felszálló repülőgépek számának eloszlását, amely egy órára vonatkozik! Legyen X a le és felszálló repülőgépek száma egy óra alatt. Mivel átlagosan 2 percenként száll le vagy fel egy gép, ezért 60 perc alatt átlagosan 30 gép száll le vagy fel.

> restart :

$$> \mu := \frac{60}{2}$$

> with(Statistics) :

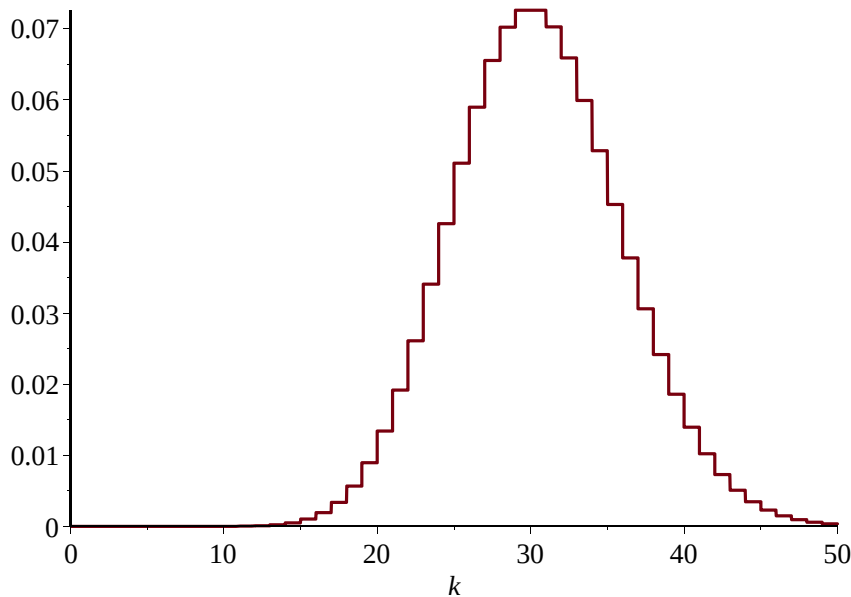
>  $X := \text{RandomVariable}(\text{Poisson}(\mu)) :$

>  $fX := \text{unapply}(\text{ProbabilityFunction}(X, k), k);$

> plot( fX(floor(k)), k = 0 ..50)

$$\mu := 30$$

$$fX := k \rightarrow \text{piecewise} \left( k < 0, 0, \frac{30^k e^{-30}}{k!} \right)$$



**(b)** Az esetek hány százalékában szükséges két irányító berendezés párhuzamos működtetése? Akkor szükséges kettőt működtetni, ha az első kapacitása betelt, azaz 20-nál több gép szeretne le- és felszállni 1 óra alatt. Nézzük meg ennek a valószínűségét a CDF függvénnyel!

>  $P(20 < X) = 1 - \text{CDF}(X, 21); \text{evalf}(\%)$

$$P(20 < \_R) = 1 - \frac{9217484362167150037}{15842827} e^{-30}$$

$$P(20 < \_R) = 0.9455565958 \quad (1.7.18)$$

Tehát az esetek 94.5%-ban szükséges még egy irányító berendezést működtetni! Azt lehet mondani, hogy szinte minden esetben legalább 2 berendezésnek kell irányítani.

**(c)** Milyen gyakran nem elegendő még két párhuzamosan működő irányító berendezés sem? Két berendezéssel 40 gép le és fel szállását lehet irányítani. Nézzük meg annak valószínűségét, hogy 40-nél több gép akar le és felszállni 1 óra alatt!

>  $P(40 < X) = 1 - \text{CDF}(X, 41); \text{evalf}(\%)$

$$P(40 < \_R) = 1 - \frac{129288200298209380523785116686765293}{12371806343732727744853} e^{-30}$$

$$P(40 < \_R) = 0.0221070400 \quad (1.7.19)$$

Tehát az esetek 2.2%-ban kellene egy harmadik irányító berendezés is. Ezért nem célszerű megvásárolni a harmadik berendezést, mert a kapacitása nem lesz kihasználva!

#### 9.7.11. PÉLDA. Binomiális és Poisson-modellek együttes alkalmazása

A következő példa tartalmazza Adams által 1936-ban közzétett mérést, amely egy útszakaszon 10 másodperces időközönként áthaladó járművek számát és ezek megfigyelt gyakoriságait mutatja. Hivatkozás: Adams, William F., "Road Traffic Considered as a Random Series," Institution of Civil Engineers Journal, Nov., 1936, pp. 121-13.

| A járművek száma,<br>amelyek egy 10 sec-os<br>periódus alatt érkeztek | A megfigyelt<br>10- sec-os<br>esetek<br>gyakorisága |
|-----------------------------------------------------------------------|-----------------------------------------------------|
| 0                                                                     | 94                                                  |
| 1                                                                     | 63                                                  |
| 2                                                                     | 21                                                  |
| 3                                                                     | 2                                                   |
| >3                                                                    | 0                                                   |

(a) Határozzuk meg a 10 másodpercenként átlagos áthaladó autók számát!  
 (b) Közelítsük X Poisson eloszlással a 10 másodperc alatt áthaladó autók számát! Rajzoljuk fel a mért és számolt gyakoriságok oszlop diagramjait!  
 (c) Közelítsük Y binomiális eloszlással a 10 másodperc alatt áthaladó autók számát!  
 Hasonlítsuk össze a mért és számolt gyakoriságok oszlop diagramjait, mind a Poisson, mind a binomiális esetben!  
 (d) Mennyi a valószínűsége, hogy 1 óra alatt több, mint 240 autó érkezik? Számoljunk Poisson-eloszlással!

### Megoldás

Az átlagosan áthaladó járművek számát megkapjuk, ha az összes autó számát elosztjuk az eltelt idővel.

> *restart*

> *jármuvekszama := [0, 1, 2, 3]*

> *gyakorisagok := [94, 63, 21, 2]*

> *osszesjarmu :=  $\sum_{k=1}^4 \text{gyakorisagok}_k \cdot \text{jarmuvekszama}_k$*

> *osszesido :=  $\sum_{k=1}^4 \text{gyakorisagok}_k$*

*jármuvekszama := [0, 1, 2, 3]*

*gyakorisagok := [94, 63, 21, 2]*

*összesjarmu := 111*

*összesido := 180*

**(1.7.20)**

Tehát Adams összesen 180 *darab* 10 *másodpercet* (= 30 perc=fél óra) mért. Ez alatt 111 gépjárművet látott elhaladni, ezért átlagosan 10 másodpercenként 111/180 jármű haladt el.

>  *$\lambda := \text{evalf}\left(\frac{\text{összesjarmu}}{\text{összesido}}\right)$*

*$\lambda := 0.6166666667$*

**(1.7.21)**

Tehát átlagosan 10 másodpercenként 0.616 gépjármű haladta el!

**(b)** Közelítsük X Poisson eloszlással a 10 másodperc alatt áthaladó autók számát! Rajzoljuk fel a mért és számolt gyakoriságok oszlop diagramjait!

Felvesszünk egy Poisson-eloszlású X változót, melynek paramétere az átlagos járműszám 10

másodperc alatt. Tudjuk, hogy a Poisson-eloszlás várható értéke a paramétere.

> *with(Statistics) :*

>  $X := \text{RandomVariable}(\text{Poisson}(\lambda)) :$

> várható érték =  $\text{Mean}(X)$

várható érték = 0.6166666667

(1.7.22)

>  $\text{modell1} := \text{összesidő} \cdot [\text{Probability}(X = 0), \text{Probability}(X = 1), \text{Probability}(X = 2), 1 - \text{Probability}(X = 0) - \text{Probability}(X = 1) - \text{Probability}(X = 2)] ;$

$\text{modell1} := [97.15330397, 59.91120412, 18.47262127, 4.462870644]$

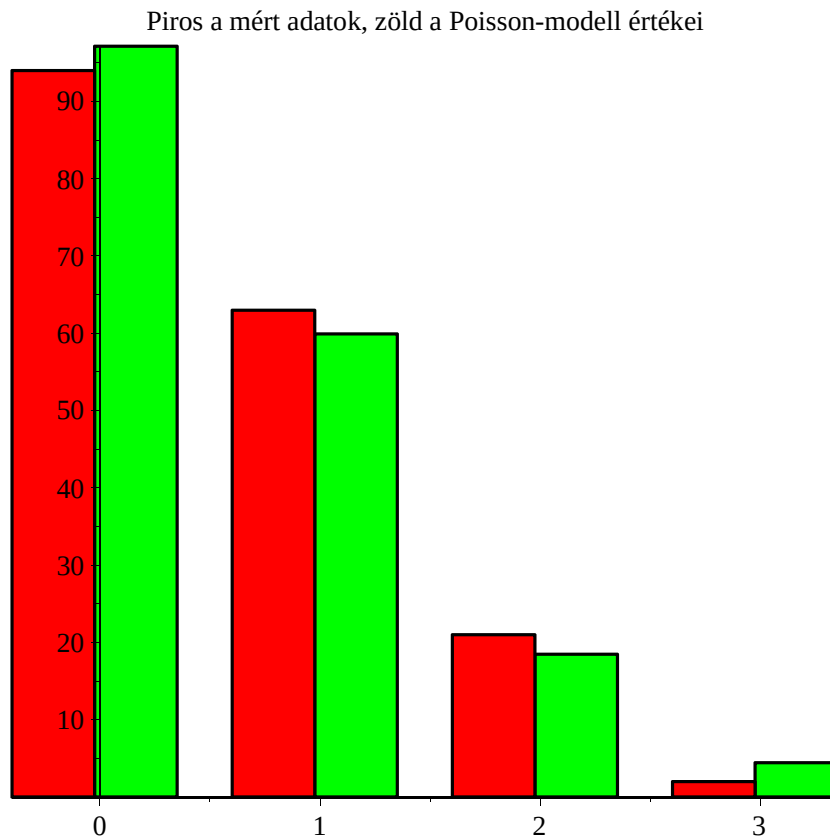
(1.7.23)

>  $\sum_{k=1}^4 \text{modell1}_k$

180.0000000000

(1.7.24)

>  $\text{ColumnGraph}([\text{gyakoriságok}, \text{modell1}], \text{offset} = -0.4, \text{color} = [\text{red}, \text{green}], \text{title} = \text{"Piros a mért adatok, zöld a Poisson-modell értékei"})$



(c) Közelítsük Y binomiális eloszlással a 10 másodperc alatt áthaladó autók számát! Hasonlítsuk össze a mért és számolt gyakoriságok oszlop diagramjait, mind a Poisson, mind a binomiális esetben!

Az Y binomiális eloszlás n és p paramétereit kell számolni úgy, hogy az

$$E(Y) = n \cdot p = E(X) = \lambda$$

Adott a  $\lambda$  és  $n$  értékét is tudjuk, mert Adams mérése során maximum 3 autót számolt egy 10 sec alatt! Ezért  $n=3$ .

>  $p := 'p'$ : *egyenlet* :=  $3 \cdot p = \lambda$

>  $\frac{\text{egyenlet}}{3}$ ; *assign*(%)

*egyenlet* :=  $3 \cdot p = 0.6166666667$

$p = 0.2055555556$  (1.7.25)

>  $Y := \text{RandomVariable}(\text{Binomial}(3, p))$

> *várható érték* = *Mean*( $Y$ )

$Y := \_R2$

*várható érték* =  $0.6166666668$  (1.7.26)

> *modell2* := *összesidő* · [*Probability*( $Y = 0$ ), *Probability*( $Y = 1$ ), *Probability*( $Y = 2$ ),  
*Probability*( $Y = 3$ )];

*modell2* := [90.25330246, 70.05675926, 18.12657407, 1.563364199] (1.7.27)

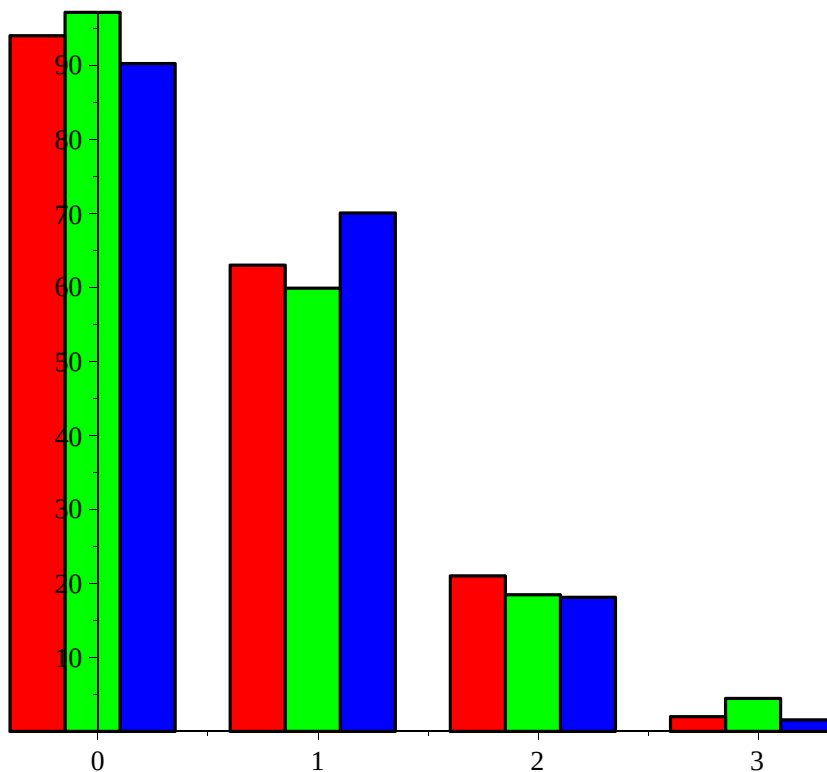
>  $\sum_{k=1}^4 \text{modell2}_k$

180.0000000000 (1.7.28)

> *ColumnGraph*( [*gyakoriságok*, *modell1*, *modell2*], *offset* = -0.4, *color* = [red, green, blue],  
*title*

= "Piros a mért adatok, zöld a Poisson-modell értékei, kék a binomiális-modell értékei")

Piros a mért adatok, zöld a Poisson-modell értékei, kék a binomiális-modell értékei



A grafikonról leolvasható, hogy mindkét elméleti eloszlás jól közelíti az autók mért eloszlását. Talán a Poisson-eloszlás jobban illeszkedik!

**(d)** Mennyi a valószínűsége, hogy 1 óra alatt több, mint 240 autó érkezik? Számoljunk Poisson-eloszlással!

Tudjuk azt, hogy független Poisson-eloszlású változók összeg Poisson-eloszlású, melynek a paramétere a két eloszlás paraméterének összege.

Nézzük meg, hogy az 1 óra hány 10 másodpercből áll!

$$> \text{tizmp} := \frac{60 \cdot 60}{10}$$

$$\text{tizmp} := 360$$

**(1.7.29)**

Tehát az 1 órát 360 darab 10 másodperces független mérésre tudjuk felosztani.

Legyen  $X_1, X_2, \dots, X_{360}$  az egyes 10 másodpercekben érkező autók véletlen száma!

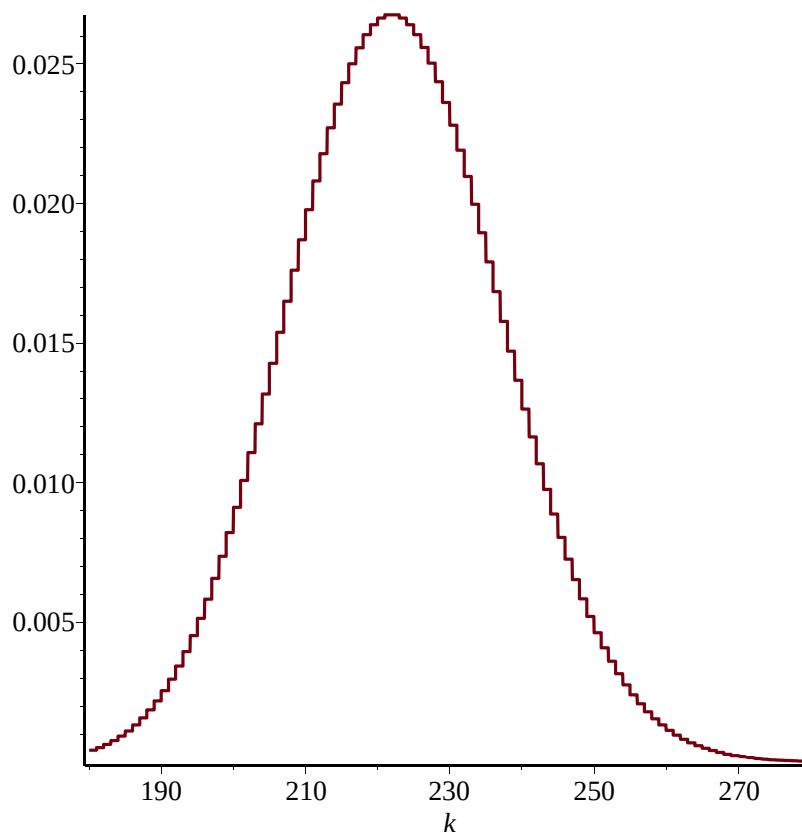
Tudjuk, hogy minden  $k$ -ra  $X_k \sim \text{Poisson}(\lambda)$ , ezért ha  $Z$  jelöli az 1 óra alatt érkező összes autók számát, akkor  $Z = X_1 + X_2 + X_3 + \dots + X_{360}$ .

A  $Z \sim \text{Poisson}(360 \cdot \lambda)$ , mert az egyes 10 mp-k alatt mért autók száma egymástól független!

**>  $Z := \text{RandomVariable}(\text{Poisson}(360 \cdot \lambda))$  :**

**>  $\text{plot}(\text{ProbabilityFunction}(Z, \text{floor}(k)), k = 180 \dots 280)$**





$$> P('Z' > 240) = 1 - CDF(Z, 240)$$

$$P(240 < Z) = 0.1082278558$$

(1.7.30)

Tehát kb. 10.8% az esély arra, hogy 1 óra alatt ilyen forgalmi adatok mellett több, mint 240 autó halad át!

#### 9.7.11. PÉLDA. Vörösvértetek gyakorisága Poisson-eloszlást követ

Egy orvosi laboratóriumban 64 vér mintát vizsgáltak a benne levő vörös vértetek száma szerint. A kapott mérési eredményeket a következő táblázat mutatja a vörös vértetek számának növekedésével és azon minták számával, melyekben a megadott számú vörös vértet volt:

|                               |   |   |   |   |    |    |   |   |    |    |    |    |    |
|-------------------------------|---|---|---|---|----|----|---|---|----|----|----|----|----|
| Vörös vértetek száma (V)      | 2 | 3 | 4 | 5 | 6  | 7  | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Hány mintában fordult elő (f) | 1 | 5 | 4 | 9 | 10 | 10 | 8 | 6 | 4  | 3  | 2  | 1  | 1  |

(a) Számítsuk ki, hogy mennyi a vörös vértetek átlagos száma egy mintában!

(b) Rajzoljuk fel a vörös vértetek relatív gyakoriságait oszlop diagrammal a Statistics csomag ColumnGraph eljárásával!

(c) Jelölje  $X$  az egy mintában előforduló vörös vértetek számát. Használjuk az  $X \sim \text{Poisson}(\lambda)$  modellt, melyben a  $\lambda$  paraméter legyen az (a) pontban számolt átlag. Rajzoljuk fel  $X$  eloszlását és tegyük a (b) pontban rajzolt oszlop diagrammal közös ábrába!

(d) Jelölje  $Y$  az egy mintában előforduló vörös vértetek számát. Használjuk az  $Y \sim \text{Binomial}(14, p)$  binomiális modellt! Számoljuk ki a  $p$  paraméter értékét úgy, hogy  $X$  és  $Y$  várható értéke legyen közös!

(e) Rajzoljuk fel közös koordináta-rendszerben  $X$  és  $Y$  eloszlását, valamint a mért relatív gyakoriságok oszlop diagramját! Döntsük el a grafikonok alapján, hogy a Poisson-modell vagy a binomiális modell illeszkedik-e jobban!

### Megoldás

```
> restart ;
> with(Statistics) :
> V := [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
> f := [1, 5, 4, 9, 10, 10, 8, 6, 4, 3, 2, 1, 1]
 V := [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
 f := [1, 5, 4, 9, 10, 10, 8, 6, 4, 3, 2, 1, 1] (1.7.31)
```

```
> N := Count(V); Count(f);
> mintak := add(f[k], k = 1..N)
 N := 13
 13
 mintak := 64 (1.7.32)
```

Tehát 13 kategória van és összesen 64 mintát vizsgáltak meg!  
Számoljuk meg, hogy összesen mennyi vörös vértetet figyeltek meg! Ehhez összegezni az  $f_k \cdot V_k$  szorzatokat!

```
> testek := add(f[k] * V[k], k = 1..N)
 testek := 450 (1.7.33)
```

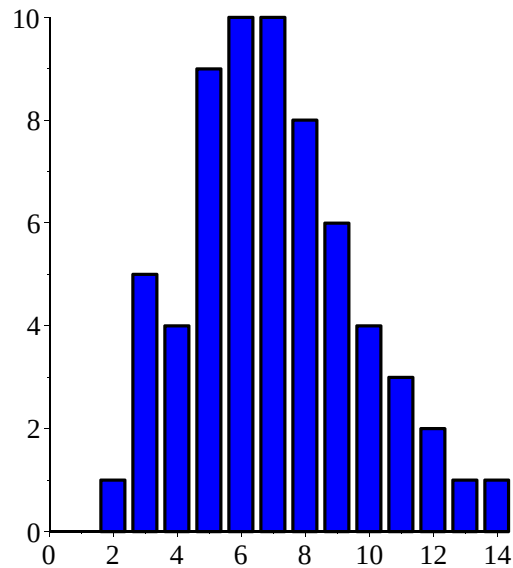
Tehát összesen 450 vörös vértetet figyeltek meg a 64 mintában! Így az egy mintában átlagosan levő vörös vértetek számát a  $\frac{\text{vörös vértetek száma}}{\text{minták száma}}$  hányados adja meg.

```
> atlag := evalf(testek / mintak)
 atlag := 7.031250000 (1.7.34)
```

Tehát kb. 7 vörös vértetet kaptak átlagosan egy mintában!

Rajzoljuk fel a gyakoriságok oszlop diagramját a ColumnGraph eljárással! Vigyázat, vörös vértetek száma nem 1-gyel, hanem kettővel kezdődik. Ezért az *offset* paraméterrel megadjuk az oszlopok jobbra tolásának mértékét!

```
> merések := ColumnGraph(f, color = blue, offset = 1.6) : merések
```

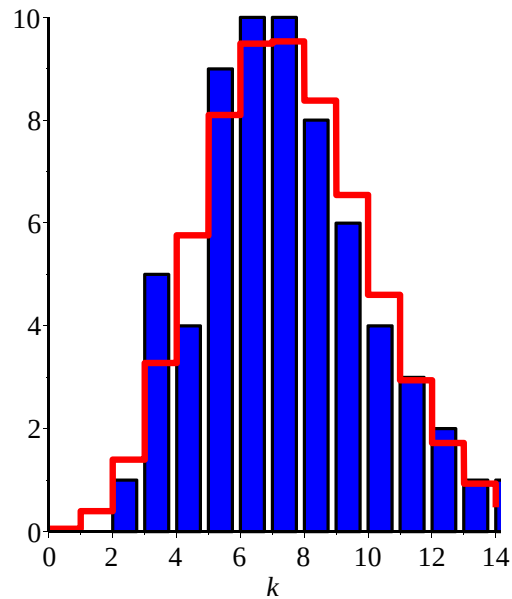


Modellezzük a vörös vértetek számának gyakoriságait  $Poisson(\lambda)$  eloszlással! Ehhez az eloszlás valószínűség értékeit megszorozzuk a minták számával, hogy a modellben számolt gyakoriságok összege is 64 legyen.

- >  $X := RandomVariable(Poisson(\text{átlag}))$
- >  $fX := unapply(mintak \cdot ProbabilityFunction(X, k), k)$
- >  $modellX := plot(fX(\text{floor}(k)), k = 0 .. 14, thickness = 3, color = red) :$
- >  $with(plots) :$
- >  $display([relativ, modellX])$

$X := \_R2$

$$fX := k \rightarrow 64 \text{ piecewise} \left( k < 0, 0, \frac{0.0008838263069 \cdot 7.031250000^k}{k!} \right)$$



Modellezzük most a vörös vértetek számának gyakoriságait  $\text{binomial}(14, p)$  eloszlással! Ehhez meg kell mondani  $p$  értékét! Az egyenlet, amiből számolunk az a várható értékek egyenlősége.

>  $p := 'p' : Y := \text{RandomVariable}(\text{Binomial}(14, p)) :$

>  $\text{egyenlet} := \text{Mean}(X) = \text{Mean}(Y);$

>  $p := \text{solve}(\text{egyenlet}, p)$

$\text{egyenlet} := 7.031250000 = 14 p$

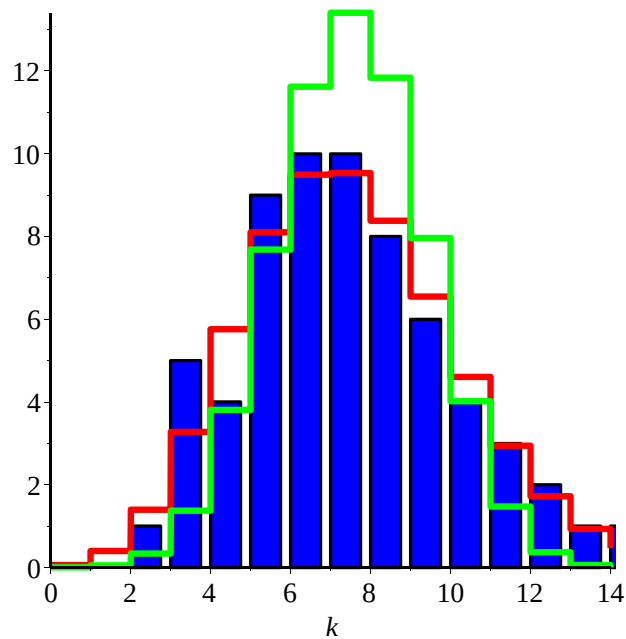
$p := 0.5022321429$

**(1.7.35)**

Most is az eloszlás valószínűség értékeit meg kell szorozni a minták számával, hogy a modellben számolt gyakoriságok összege 64 legyen.

>  $\text{modellY} := \text{plot}(\text{mintak} \cdot \text{ProbabilityFunction}(Y, \text{floor}(k)), k = 0 .. 14, \text{thickness} = 3, \text{color} = \text{green}) :$

>  $\text{display}([\text{relativ}, \text{modellX}, \text{modellY}])$



Láthatóan a Poisson-eloszlás jobban illeszkedik a mért értékekhez, mint a binomiális -eloszlású értékek!

## ▼ 9.8. Geometriai eloszlás

Tekintsük újból a **Bernoulli-féle kísérletet**:

- (a) Legyen egy kísérlet kimenetele **sikeres** vagy **sikertelen** az A esemény szempontjából.
- (b) Ismételjük egymásután többször a kísérletet **függetlenül**.
- (c) A kísérlet egy végrehajtása során a **sikeres** kimenetel valószínűsége legyen  $p = P(A)$ . ( $0 < p < 1$ )

### 9.8.1. DEFINÍCIÓ. Geometriai eloszlás

A Bernoulli-kísérlet során jelölje  $X$  a kísérlet **első sikeres kimenetelének sorszámát**, vagyis  $X$  jelöli azt a sorszámot, ahányadikra a kísérlet először sikeres! Az  $X$  valószínűségi változót **geometriai eloszlásúnak** nevezzük és  $p$  a paramétere:  $X \sim \text{Geometriai}(p)$

$X$  felvehet tetszőleges pozitív egész értéket

$$X = 1, 2, 3, \dots$$

mely azt jelenti, hogy lehet a kísérlet kimenetele sikeres elsőre, vagy ha elsőre nem, akkor másodikra, vagy ha elsőre és másodikra nem, akkor harmadikra,... stb.

Ennek megfelelően a valószínűségek

$$\begin{aligned} P(X=1) &= P(A) = p, \\ P(X=2) &= P(\bar{A} \cdot A) = P(\bar{A}) \cdot P(A) = (1-p) \cdot p, \\ P(X=3) &= P(\bar{A} \cdot \bar{A} \cdot A) = P(\bar{A}) \cdot P(\bar{A}) \cdot P(A) = (1-p)^2 p \\ P(X=k) &= (1-p)^{k-1} p. \end{aligned}$$

Nevét az eloszlás onnan kapta, hogy a fenti valószínűségek végtelen sorozata egy mértani vagy geometriai sorozatot alkot, melyben a képzési szabály rekurzívan úgy adható meg, hogy egy sorozat elem az előző sorozat elem  $q = (1-p)$ - szerese.

### 9.8.2. TÉTEL. A geometriai eloszlás érték táblázata

Ha  $X \sim \text{Geometriai}(p)$ , akkor

$$P(X = n) = (1 - p)^{n-1} \cdot p$$

az eloszlás értékei  $n = 1, 2, 3, \dots$  egészekre.

#### Bizonyítás

Mutassuk meg, hogy a megadott valószínűségek összege 1.

$$\sum_{n=1}^{\infty} (1-p)^{n-1} \cdot p = p \cdot \sum_{n=0}^{\infty} (1-p)^n = p \cdot \frac{1}{1 - (1-p)} = p \cdot \frac{1}{p} = 1,$$

mert a kapott mértani sor konvergens, mivel a  $q = 1 - p$  kvóciense 0 és 1 közé esik. Q.e.d.

Vegyünk fel Maple-ben egy geometriai eloszlású  $X$  diszkrét valószínűségi változót  $p$  paraméterrel!

> *restart : with(Statistics) :*

> *X := RandomVariable(Geometric(p)) : f := unapply(ProbabilityFunction(X, k), k) :  
P('X' = k) = f(k)*

$$P(X = k) = \begin{cases} 0 & k < 0 \\ p (1 - p)^k & \text{otherwise} \end{cases} \quad (1.8.1)$$

**Figyelem**, látható, hogy a Maple-ben felvett geometriai eloszlás értéke  $X=0$ -tól indul és nem  $X=1$ -től. Ez az úgy nevezett **módosított geometriai eloszlás**, amely a **sikertelen kísérletek számát adja meg az első sikeres kísérlet előtt**.

Ha  $X$  helyett az  $Y=X+1$  valószínűségi változót vesszük fel, akkor  $Y$  értéke 1 -től indul és megegyezik a fennnt értelmezett geometriai eloszlással! Nézzük a Maple meg tudja-e adni az  $(X+1)$  transzformációhoz tartozó diszkrét eloszlást?

> *Probability(X + 1 = k) ;*

$$\begin{cases} 0 & k < 1 \\ p (1 - p)^{k-1} & \text{otherwise} \end{cases} \quad (1.8.2)$$

Rajzoljuk fel  $X$  eloszlását és az **eloszlás függvényét**, ha  $p=0.4$ !

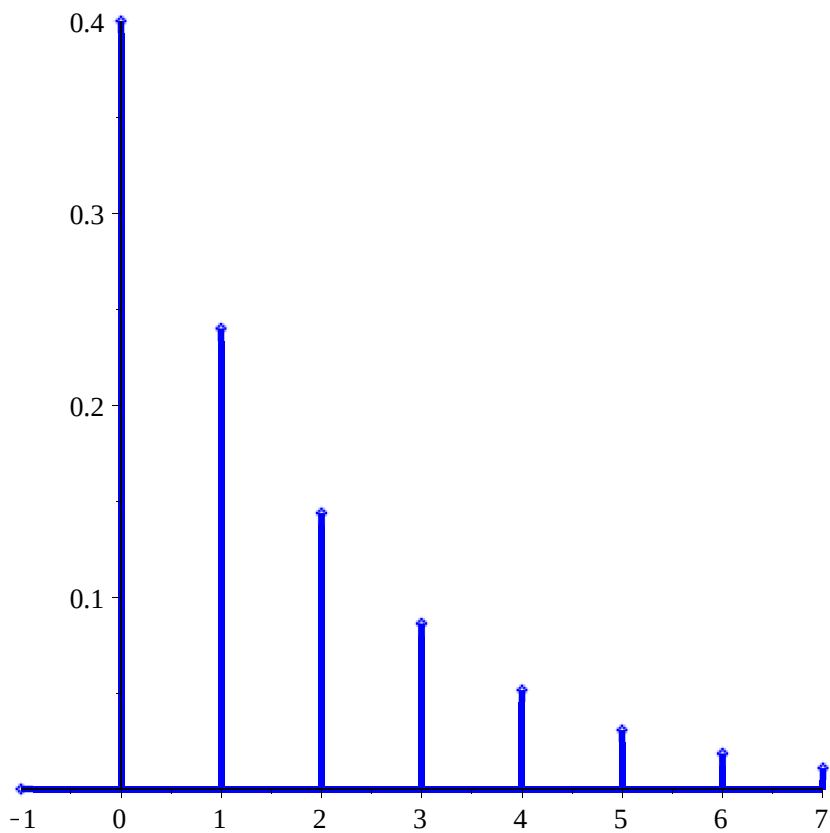
> *XX := [seq(x, x = -1 .. 7)] ;*

*YY := eval(map(f, XX), p = 0.4)*

*XX := [-1, 0, 1, 2, 3, 4, 5, 6, 7]*

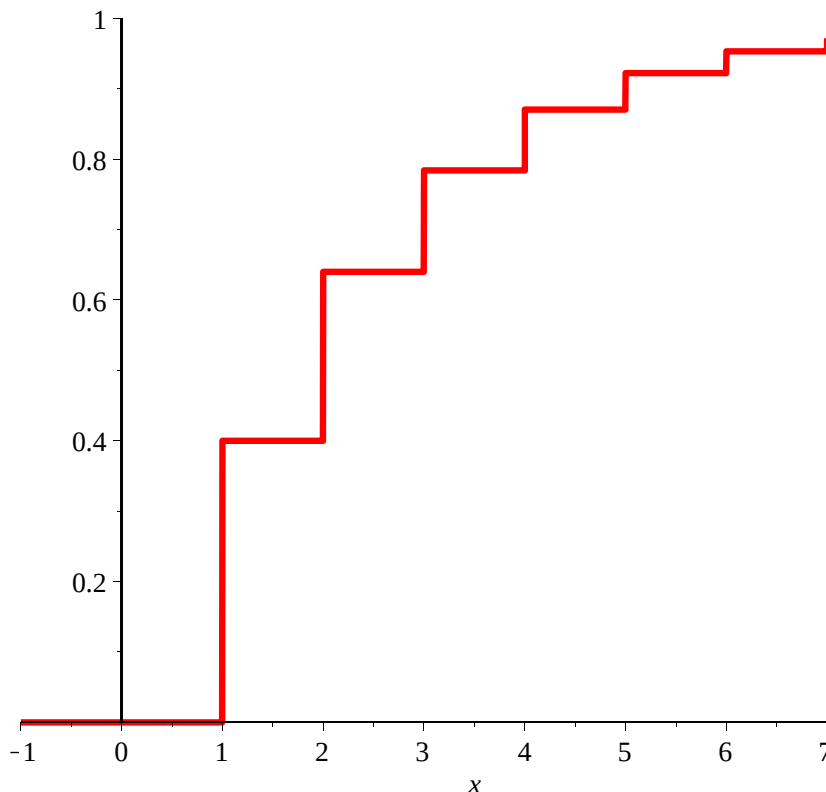
*YY := [0, 0.4, 0.24, 0.144, 0.0864, 0.05184, 0.031104, 0.0186624, 0.01119744]* (1.8.3)

> *DynamicSystems:-DiscretePlot(XX, YY, style = stem, thickness = 3, color = blue)*



➤ *plot(CDF(Geometric(0.4), x - 1), x = -1 ..7, 0 ..1, thickness = 2, color = red, title = Geometriai eloszlásfüggvény, ha p=0.4, thickness = 3)*

Geometriai eloszlásfüggvény, ha  $p=0.4$



### 9.8.2. TÉTEL. Geometriai eloszlás várható értéke és varianciája

Az  $X \sim \text{Geometriai}(p)$  eloszlás **várhatóértéke**  $E(X) = \frac{1}{p}$  és **varianciája**  $\text{Var}(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$ .

#### Bizonyítás

A diszkrét eloszlás várható értékének számítási szabálya alapján

$$E(X) = \sum_{k=1}^{\infty} k \cdot q^{k-1} \cdot p = \sum_{k=1}^{\infty} \frac{d q^k}{d q} \cdot p = p \cdot \frac{d}{d q} \sum_{k=1}^{\infty} q^k = p \cdot \frac{d}{d q} \left( \frac{q}{1-q} \right) = p \cdot \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p},$$

ahol kihasználtuk, hogy a deriválás és a végtelen összegzés felcserélhető a geometriai sornál a konvergencia tartományon belül.

Ugyanez Maple-ben.

$$> \sum_{k=1}^{\infty} k p (1-p)^{k-1} = \sum_{k=1}^{\infty} k p (1-p)^{k-1};$$

Várható érték a  $\text{Mean}(X+1)$  hívásával =  $\text{simplify}(\text{Mean}(X+1))$

$$\sum_{k=1}^{\infty} k p (1-p)^{k-1} = \frac{1}{p}$$



$$\text{Várható érték a } \text{Mean}(X+1) \text{ hívásával} = \frac{1}{p} \quad (1.8.4)$$

Hasonlóan megy, de kicsit több számolással a variancia levezetése. Nézzük ezt csak Maple-ben.

$$> \sum_{k=1}^{\infty} k^2 p (1-p)^{k-1} - \text{Mean}(X)^2 = \sum_{k=1}^{\infty} k^2 p (1-p)^{k-1} - \frac{1}{p^2};$$

A variancia összegzéssel = simplify(rhs(%));

> A variancia a Variance eljárással = Variance(X)

$$\sum_{k=1}^{\infty} k^2 p (1-p)^{k-1} - \frac{(1-p)^2}{p^2} = \frac{2-p}{p^2} - \frac{1}{p^2}$$

$$\text{A variancia összegzéssel} = -\frac{-1+p}{p^2}$$

$$\text{A variancia a Variance eljárással} = \frac{1-p}{p^2} \quad (1.8.5)$$

Mint láttuk korábban, az eltolás nem változtatja meg a variáciát. Tehát  $\text{Var}(X) = \text{Var}(X+1)$ .

> Variancia a Variance(X+1) hívással = Variance(X+1); simplify(%)

$$\text{Variancia a Variance}(X+1) \text{ hívással} = \frac{1-p}{p^2} + \left( \frac{1-p}{p} + \frac{-1+p}{p} \right)^2$$

$$\text{Variancia a Variance}(X+1) \text{ hívással} = -\frac{-1+p}{p^2} \quad (1.8.6)$$

### 9.8.3. DEFINÍCIÓ. VALÓSZÍNŰSÉGI VÁLTOZÓ MEMÓRIA NÉLKÜLI TULAJDONSÁGA.

Az alábbi két definíció egyenértékű, amelyek az  $X$  eloszlás **memória nélküli tulajdonságát** fejezik ki.

$$(a) \quad P(X > k+m \mid X > m) = P(X > k)$$

vagy ezzel ekvivalens módon

$$(b) \quad P(X-m < k \mid X > m) = P(X < k) = F_X(k)$$

Számítsuk ki az  $X \sim \text{Geometriai}(p)$  eloszlás  $P(X \leq m)$  valószínűségét

$$P(X \leq m) = \sum_{k=1}^m p \cdot q^{k-1} = p \cdot (1 + q + \dots + q^{m-1}) = (1-q) \cdot (1 + q + \dots + q^{m-1}) = 1 - q^m,$$

ahol  $m = 1, 2, \dots$  lehet.

Számítsuk ki az  $F_X(m) = P(X < m)$  eloszlásfüggvény értékét is

$$F_X(X) = P(X < m) = P(X \leq m) - P(X = m) = 1 - q^m - p \cdot q^{m-1} = 1 - q^{m-1}(q + p) = 1 - q^{m-1}$$

Ennyi előkészület után igazoljuk a geometriai eloszlás memória nélküli tulajdonságát.

**9.8.4. TÉTEL. A geometriai eloszlás rendelkezik az emlékezet nélküli tulajdonsággal**  
Tetszőleges  $X \sim \text{Geometriai}(p)$  eloszlás rendelkezik a **memória nélküli tulajdonsággal**.

#### Bizonyítás

(a) tulajdonság bizonyítása

Mivel  $P(X \leq m) = 1 - q^m$  tetszőleges  $m=1, 2, 3, \dots$  esetén. Ahonnan

$$P(X > m) = 1 - P(X \leq m) = 1 - (1 - q^m) = q^m = (1 - p)^m.$$

A memória nélküli tulajdonság (a) egyenlőségének baloldalát alakítva eljutunk az egyenlet jobb oldalához!

$$\begin{aligned} P(X > k + m | X > m) &= \frac{P(k + m < X, m < X)}{P(X > m)} = \frac{P(X > k + m)}{P(X > m)} = \\ &= \frac{(1 - p)^{k+m}}{(1 - p)^m} = (1 - p)^k = P(X > k). \end{aligned}$$

**(b)** Nézzük a memória nélküli tulajdonság (b) alakjának igazolását, amelyben az eloszlás függvény szerepel!

A feltételes valószínűség definíciója alapján

$$\begin{aligned} P(X - m < k | X > m) &= \frac{P(m < X < k + m)}{P(X > m)} = \frac{F_X(k + m) - F_X(m + 1)}{(1 - p)^d} = \\ &= \frac{1 - q^{k+m-1} - (1 - q^m)}{(1 - p)^m} = \frac{q^m - q^{k+m-1}}{q^m} = 1 - q^{k-1} = P(X < k) = F(k) \end{aligned}$$

Tehát az  $Y = X - d$  "maradék" változó eloszlás függvénye ugyanaz, mint az eredeti  $X$  változó eloszlás függvénye, feltéve, hogy  $X > d$ .

**Megjegyezzük**, hogy a diszkrét eloszlások között a geometriai eloszlás az egyedüli, amely rendelkezik a memória nélküli tulajdonsággal.

#### 9.8.5. PÉLDA. Sorsjegy húzás első nyerésig

Kapáros sorsjegy szerencsejátékhoz elkészítettek  $n=10\,000$  darab sorsjegyet. A sorsjegyek között 10 darab egyenként 100 000 Ft-os díj van elrejtve.

**(a)** Várhatóan hány sorsjegyet adnak el, mire megtalálja valaki az első nyereményt tartalmazó szelvényt!

**(b)** Mennyiért kellene árulni egy sorsjegyet, ha az első nyereményig befolyt összeg duplája kellene, hogy legyen az elvitt 100 000,-Ft-os nyereménynek?

#### Megoldás

A megtalálás valószínűsége minden alkalommal  $p = \frac{10}{10000} = 0.001$ , mert ha 1-t kivesszünk a sokaságból, akkor a nyerés esélye jelentősen nem növekszik a sorsjegyek nagy száma miatt. Jelölje  $X$  azt a sorszámot, ahányadikra megtalálják az első nyertes sorsjegyet. Ekkor  $X \sim \text{Geometria}(p)$  és ennek várható értékét  $E(X) = \frac{1}{p}$  kérdezi a feladat (a) része.

> restart : with(Statistics) :  $p := \frac{10}{10000}$

$$p := \frac{1}{1000} \quad (1.8.7)$$

>  $X := \text{RandomVariable}(\text{Geometric}(p)) :$

>  $f := \text{unapply}(\text{ProbabilityFunction}(X, k), k) : f(k)$

$$\begin{cases} 0 & k < 0 \\ \frac{1}{1000} \left( \frac{999}{1000} \right)^k & \text{otherwise} \end{cases} \quad (1.8.8)$$

> Mean(X + 1)

$$1000 \quad (1.8.9)$$

Tehát várhatóan 1000 darabot adnak el a 10 000-ből, mire valaki megtalálja az első nyereményt tartalmazó szelvényt! Ez a tény magától értetődő, mert 1000 darabonként átlagosan van egy nyeremény. Hiszen a 10 nyeremény 10 000 szelvény között van egyenletesen szétterítve.

**(b)** Mennyiért kellene árulni egy sorsjegyet, ha az első nyereményig befolyt összeg duplája kellene, hogy legyen az elvitt 100 000,-Ft-os nyereménynek?

Mivel a nyereményre ki kell fizetni 100 000 Ft.-ot és ennek kétszeresét 200 000, Ft-ot szeretnének beszédni, továbbá várhatóan 1000 darabot adnak el az első nyereményig, ezért a 200 000, Ft bevételt 1000 egyenlő részre kell szétosztani!

$$> 2 \cdot 100000 / 1000;$$

200

(1.8.10)

Tehát egy sorsjegyet legalább 200,- Ft-ért kell árulni. ♣

### 9.8.6. PÉLDA. Cache-memória elérése

A **cache** (ejtsd kes) memória a számítógép **gyors elérésű** és általában 1 MB alatti közvetlen elérésű tár területe. Célja a memóriához fordulási idő csökkentése. A cache memória **másolatot tárol** a megelőzően használt adatokból, programokból. Ha a felhasználó kéri valamely adatsor **betöltését**, akkor a számítógép először **ellenőrzi**, hogy a cache memóriában megtalálható-e a kért adatról másolat. Ha megtalálja a cache-ben, akkor azt gyorsabban tölti be, mint ha a memóriából töltötte volna be. Ha nem találja a cache-ben csak, akkor fordul a memóriához, hogy a kért adatot onnan töltsse be. Ez utóbbi betöltés lassúbb, mint ha a cache-ből töltötte volna be. Majd gondoskodik arról, hogy a cache-ben régóta tárolt adatrész törölődjön és az új adat kerüljön helyére.

Tegyük fel, hogy egy kért adatsor a cache-ben  $p$  valószínűséggel található meg. Határozzuk meg annak valószínűségét, hogy az egymás utáni cache memóriához fordulási kérelmek közül az **első sikertelen hozzáférés** a  $k$  – ik kérés lesz. Feltesszük, hogy a kért adat jelenléte a cache-ben kérelmenként egymástól **független**.

### Megoldás

Legyen  $X_i = 1$ , ha az  $i$ -edik cache-hez fordulási kérelem sikeres és egyébként  $X_i = 0$ . A feltétel szerint  $P(X_i = 1) = p$  és  $P(X_i = 0) = 1 - p$ , továbbá az  $X_1, X_2, \dots$  Bernoulli -változók sorozata egy független sorozat.

A keresett eseményt jelölje  $T = k$ , amely azt jelenti, hogy az **első sikertelen cache-hez fordulási kérelem** a  $k$ -adik hozzáfutási kérelem. Akkor lesz a  $k$ -adik az első sikertelen kérelem, ha az előtte levő  $(k-1)$  kérelem sikeres volt. Tehát felírhatjuk a

$$\{T = k\} = \{X_1 = 1\} \cdot \{X_2 = 1\} \cdot \dots \cdot \{X_{k-1} = 1\} \cdot \{X_k = 0\}$$

események egyenlőségét.

A függetlenség alapján  $P(T = k) = p^{k-1} (1 - p)$ , ahol  $k = 1, 2, 3, \dots$ . Tehát a  $T$  változó geometriai eloszlású. ♣

### 9.8.7. PÉLDA. Hulló csillagok

Május és október hónapok között láthatunk hulló csillagokat átlagosan 20 percenként egyet. Ha minden nap este 1 órahosszat ülünk kint és várjuk a hulló csillag megjelenését, akkor hány napig kell átlagosan figyelni, hogy 10 vagy 10-nél több hulló csillagot figyeljünk meg egyetlen napon.

### Megoldás

A példa érdekes módon ötvözi a Poisson és a Geometriai eloszlást.

Jelölje  $N$  a hullócsillagok számát, amelyet megfigyelünk egy nap alatt, amikor kiülünk 1 órára. Ekkor az

$$A = \{N \geq 10\}$$

eseményt szeretnénk elérni és jelölje  $p = P(A)$  ennek a valószínűségét. Az  $N$  változó modellezhető  $N \sim \text{Poisson}(\lambda)$ , ahol  $\lambda = 3$

Poisson-eloszlással. Ugyanis az 1 óra megfigyelési idő 3 darab 20 percből áll és 20 percenként átlagosan 1 hullócsillagot lehet látni. A Poisson-eloszlás számítási képlete alapján

$$p = P(N \geq 10) = \sum_{k=10}^{\infty} \frac{3^k}{k!} e^{-3}.$$

Ennek közelítő értékét számítsuk ki Maple-vel!

> *restart : with(Statistics) :*  
*N := RandomVariable(Poisson(3)) :*  
*Probability(N ≥ 10); p := evalf(%)*

$$1 - \frac{22471}{1120} e^{-3}$$

$$p := 0.0011024879 \quad (1.8.11)$$

Tehát kb.  $p = 0.0011$  annak a valószínűsége, hogy egy nap 1 órát figyelve 10 vagy 10-nél több hullócsillagot pillantunk meg. Erre az A eseményre végzünk független kísérleteket egymás után. Tehát klasszikus Bernoulli-kísérletet végzünk. Azt kérdezzük, hogy mikor lesz az első nap, amikor az A esemény bekövetkezik?

Jelölje  $X$  azon napok számát, amennyi eltelik addig, amíg először fordul elő az A esemény vagyis, hogy 10 vagy 10-nél több hullócsillagot figyelünk meg egy nap. Ekkor  $X \sim \text{Geometric}(p)$  eloszlású, azzal a  $p$  valószínűséggel, amelyet a Poisson-eloszlásból kaptunk. A kérdés az  $X$  változó várható értéke, amely  $E(X) = \frac{1}{p}$ .

> *X := RandomVariable(Geometric(p)) :*  
*napok száma átlagosan = Mean(X + 1)*

$$\text{napok száma átlagosan} = 907.0394329000$$

$$(1.8.12)$$

Tehát átlagosan 907 napig kell naponta egy órát figyelni, hogy a kívánt esemény eljőjjön. Ez több, mint 30 hónap. Mivel az év májusa és októbere közötti 6 hónapban lehet csak megfigyelni hullócsillagot, ezért ez a várakozás kb. 5 évet vesz igénybe.

### 9.8.8. A geometriai-eloszlás véges változata

Van  $n$  darab villanykörte, melyek között van néhány selejtes. Csak úgy tudjuk megmondani, hogy melyik jó és melyik selejtes, ha becsavarjuk az izzókat egyesével egy foglalatba és felkapcsoljuk az áramot. Tudjuk, hogy egy lámpa  $p$  valószínűséggel selejtes ( $0 < p < 1$ ). Addig vizsgáljuk át a lámpákat egyesével (és a rosszat eldobjuk), amíg egy jót nem találunk. Jelölje  $X$  az átvizsgált lámpák számát, mire megtaláltuk az első jót!

(a) Készítsük el  $X$  érték- és valószínűség táblázatát  $n=10$  esetén! Rajzoljuk fel az eloszlás pálcika diagramját!

(b) Számoljuk ki, hogy várhatóan mennyi lámpát kell átvizsgálni tetszőleges  $n$  és  $p$  esetén!

(c) Rajzoljuk fel a várható értéket  $n=100$  esetén  $p$  függvényében!

(d) Mekkora  $p$  valószínűség mellett lesz  $X$  várható értéke 2, ha  $n=100$ ?

### MEGOLDÁS

(a) Legalább egy lámpát át kell vizsgálni, hogy találjunk egy jót, ezért  $X$  legkisebb értéke 1.  $X=1$  esetén elsőre jót találunk, melynek valószínűsége  $= (1 - p)$

$X=2$  esetén elsőre rosszat találtunk, melynek valószínűsége  $p$  és másodikra jót, ennek valószínűsége  $(1 - p)$ . A két esemény együttes bekövetkezése a függetlenség miatt a két esemény valószínűségének szorzata. Tehát  $P(X=2) = p (1 - p)$ . Hasonló gondolatmenettel

kapjuk, hogy  $P(X=k) = p^{k-1} (1 - p)$ , amennyiben  $k=1, 2, 3, 4, 5, 6, 7, 8, 9$ . Az  $X=10$  esemény valószínűségét ettől eltérően kell számolni, mert csak 10 darab izzónk van! Ezért akár találunk jó

izzót, akár nem  $X=10$  lesz az átvizsgált lámpák száma.

$P(X=10) = P(\{\text{az első kilenc selejtes és a 10-ik jó}\} \text{ vagy } \{\text{az első kilenc selejtes és a 10-ik is selejtes}\}) = p^{10} + p^9 (1-p) = p^9 (p + 1 - p) = p^9$ .

> *restart : n := 10; X := [seq(k, k = 1..10)]*

> *PX := [seq( $p^{k-1} (1-p)$ ,  $k = 1..n-1$ ),  $p^{n-1}$ ]  
n := 10*

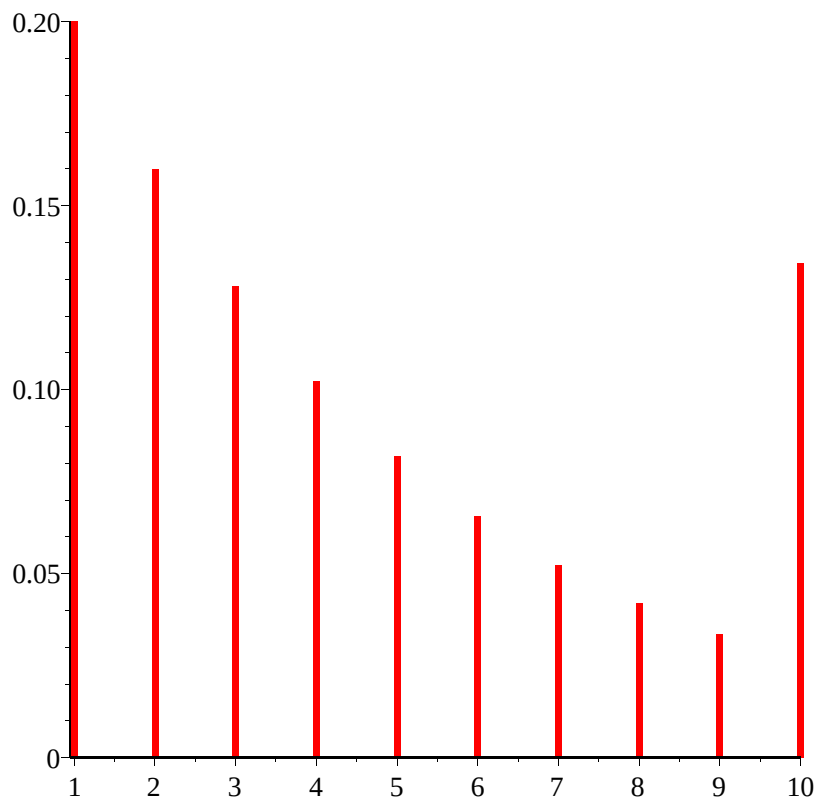
*X := [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]*

*PX := [1-p, p(1-p),  $p^2(1-p)$ ,  $p^3(1-p)$ ,  $p^4(1-p)$ ,  $p^5(1-p)$ ,  $p^6(1-p)$ ,  $p^7(1-p)$ ,  $p^8(1-p)$ ,  $p^9$ ]* **(1.8.13)**

Az eloszlás felrajzolásához a pálcikák végpontjai koordinátáinak sorozatát kell megadni.

*palcika := subs(p = 0.8, [seq([ $X_k$  0], [ $X_k$   $PX_k$ ]),  $k = 1..n$ ]) :*

> *plot(palcika, color = red, thickness = 3)*



Nézzük meg, hogy eloszlást kaptunk-e?

$$\sum_{k=1}^n PX_k = \text{simplify} \left( \sum_{k=1}^n PX_k \right)$$

$$1-p + p(1-p) + p^2(1-p) + p^3(1-p) + p^4(1-p) + p^5(1-p) + p^6(1-p) \quad \text{(1.8.14)}$$

$$+ p^7 (1 - p) + p^8 (1 - p) + p^9 = 1$$

Így valóban diszkrét valószínűségi eloszlást kaptunk, mert a valószínűségek összege 1.

Az  $X$  változó eloszlása hasonlít a geometriai eloszlásra azzal a kivétellel, hogy a geometriai eloszlás tetszőleges természetes számra értelmezett, míg a most megadott valószínűségi változó  $n$ -nél nagyobb értéket nem vehet fel!

**(b)** Számoljuk ki, hogy várhatóan mennyi lámpát kell átvizsgálni tetszőleges  $n$  és  $p$  esetén!

Ehhez  $X$  várható értékét kell meghatározni  $n$  és  $p$  függvényében!

$$> n := 'n': \sum_{k=1}^{n-1} k p^{k-1} (1 - p) + n p^{n-1}; \mu := \text{simplify}(\%)$$

$$- \frac{p^n (-n + n p - p)}{p (-1 + p)} - \frac{1}{-1 + p} + n p^{n-1}$$

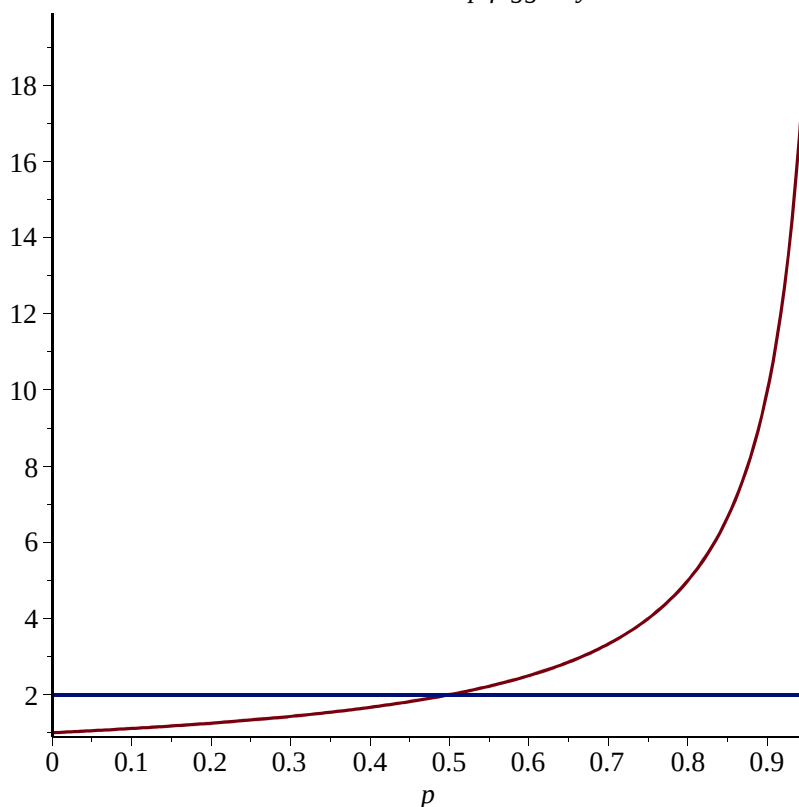
$$\mu := \frac{p^n - 1}{-1 + p} \quad (1.8.15)$$

**(c)** Rajzoljuk fel a várható értéket  $n=100$  esetén  $p$  függvényében!

$p := 'p':$

$\text{plot}([ \text{subs}(n = 100, \mu), 2 ], p = 0 .. 0.95, \text{title} = \text{A várható érték növekedése } p \text{ függvényében})$

A várható érték növekedése  $p$  függvényében



**(d)**  $E(X)=2$  egyenlet megoldása  $p$ -re, ha  $n=100$ .

$fsolve(subs(n = 100, \mu) = 2, p)$

0.5000000000

(1.8.16)

Tehát, akkor kell 100 lámpa esetén átlagosan 2-t átvizsgálni, ha tetszőleges lámpa 50% ban lehet selejtes illetve 50% -ban lehet nem selejtes.

## ▼ 9.9. Elméleti ellenőrző kérdések

- 9.9.1. Hogyan értelmezzük a diszkrét egyenletes eloszlást?
- 9.9.2. Mennyi a várható értéke és varianciája a diszkrét egyenletes eloszlásnak?
- 9.9.3. Hány paramétere van Bernoulli eloszlásnak?
- 9.9.4. Mennyi a várható értéke és varianciája a Bernoulli- eloszlásnak?
- 9.9.5. Milyen kísérletet nevezünk Bernoulli -kísérletnek?
- 9.9.6. Hogyan értelmezzük a binomiális eloszlást?
- 9.9.7. Milyen kapcsolat van a binomiális eloszlás és a Bernoulli-eloszlás között?
- 9.9.8. Mennyi a várható értéke és varianciája a binomiális- eloszlásnak?
- 9.9.9. Mik lesznek a binomiális eloszlás paraméterei a visszatevéses mintavétel esetén?
- 9.9.10. Mit a jelentése az  $N$ ,  $S$ ,  $n$  paramétereknek a hipergeometrikus eloszlásban?
- 9.9.11. Milyen képlettel értelmezzük a hipergeometrikus eloszlást?
- 9.9.12. Mennyi a várható értéke és varianciája a hipergeometrikus- eloszlásnak?
- 9.9.13. Milyen képlettel értelmezzük a Poisson- eloszlást?
- 9.9.12. Mennyi a várható értéke és varianciája a Poisson- eloszlásnak?
- 9.9.14. Milyen kapcsolat van a binomiális és Poisson-eloszlások között?
- 9.9.15. Milyen eloszlású a  $Poisson(\lambda)$  és a  $Poisson(\mu)$  független eloszlások összege?
- 9.9.16. Milyen képlettel értelmezzük a geometriai- eloszlást?
- 9.9.17. Mennyi a várható értéke és varianciája a geometriai- eloszlásnak?
- 9.9.18. Mit jelent a geometriai eloszlás memória nélküli tulajdonsága?

## ▼ 9.10. Gyakorló feladatok

### 9.10.1. feladat

Egy csomagküldő cég folyamatosan kap megrendeléseket telefonon keresztül. A menedzser szeretné tudni a megrendelések eloszlását, ezért rögzíti az egy nap alatt érkezett megrendelések számát. Egy nap alatt kapott megrendelések száma 0 és 5 között változott az alábbi gyakoriság értékekkel.

*Egy nap alatt a megrendelések száma*  $:= [0, 1, 2, 3, 4, 5]$

*Napok gyakorisága, ahányszor előfordultak a fenti megrendelési számok*  $:= [8, 13, 10, 6, 2, 1]$

- (a) Számítsa ki, hogy naponta átlagosan hány megrendelés érkezett! ( $\lambda = ?$ )
- (b) Közelítse a napok gyakoriság értékeit egy Poisson-eloszlású  $X$  változóval, ahol  $X$  jelöli az egy nap alatti megrendelések számát!
- (c) Közelítse a napok gyakoriság értékeit egy Binomiális-eloszlású  $Y$  változóval, ahol  $Y$  jelöli az egy nap alatti megrendelések számát!
- (d) Rajzolja fel a mért napok gyakoriságait, a Poisson-modell illetve binomiális modell alapján számolt gyakoriságok oszlop diagrammjaikat közös koordináta-rendszerben! Melyik modell ad jobb közelítést?
- (e) Minimálisan mennyi lesz a napi megrendelések azon  $K$  száma, amelynél több megrendelés valószínűsége már 0.1-nél kisebb? Számoljon a Poisson-eloszlással!
- (f) Számolja az (e) feladatot Markov -egyenlőtlenséggel! Összhangban vannak-e az (e) és (f) feladatok megoldásai?

### 9.10.2. feladat

Egy doboz  $N=75$  darab tömítő gyűrűt tartalmaz, amelyből 5 darabnak más a vastagsága, ezért nem elfogadható minőségű. A 75 darab tömítő gyűrűből véletlenszerűen kivesszünk  $n=10$  gyűrűt. Jelölje  $X$  a 10 kivett darabban található nem elfogadható gyűrűk számát!

- (a) Adjuk meg  $X$  eloszlását, ha a kivett gyűrűket nem tesszük vissza!
- (b) Adjuk meg  $X$  eloszlását, ha a kivett gyűrűket megvizsgáljuk, majd visszatesszük!
- (c) Rajzoljuk fel az (a) és (b) esetben kapott eloszlásokat!
- (d) Határozzuk meg a várható értéket és szórást mindkét esetben!
- (e) Hasonlítsuk össze a várható értékeket! Hasonlítsuk össze a szórásokat!

### 9.10.3. feladat

Egy népszerű Web oldal látogatottsága **Poisson** eloszlást követ naponta átlagosan 1000 látogatóval. Határozzuk meg a következőket!

- (a) Rajzoljuk fel a látogatottság eloszlását és eloszlásfüggvényét!
- (b) Mekkora a valószínűsége, hogy egy nap 1050-nél többen látogatják a Web oldalt?
- (c) Mekkora a valószínűsége, hogy egy nap 940-nél kevesebben látogatják a Web oldalt?
- (d) Mennyi a látogatók minimális száma naponta, ha a napi látogatottság valószínűsége meghaladja a  $p=0.05$  értéket?
- (e) Határozzuk meg annak valószínűségét, hogy egy év (365 nap) alatt 15-nél több olyan nap lesz, amelyeken a látogatók száma 1010-nél több!

### 9.10.4. feladat

Nyomatott áramköröket tesztelünk működés szempontjából. Egy sokaság 140 áramkört tartalmaz, melyből 10 áramkör selejtes.

A 140 áramkör közül véletlenszerűen kivesszük  $n = 15$  darabot.

Jelölje  $X$  a 15 kivett darabban található selejtes áramkörök számát!

- (a) Adjuk meg  $X$  eloszlását, ha a kivett darabokat nem tesszük vissza!
- (b) Adjuk meg  $X$  eloszlását, ha a kivett darabokat megvizsgáljuk, majd visszatesszük!
- (c) Rajzoljuk fel az (a) és (b) esetben kapott eloszlásokat!
- (d) Határozzuk meg a várható értéket és szórást mindkét esetben!
- (e) Hasonlítsuk össze a várható értékeket! Hasonlítsuk össze a szórásokat!

### 9.10.5. feladat

Az olaj keresés során az Északi - tengerben átlagosan minden 500 kutató fúrásból csak 1 fúrás esetén találnak olajat.

(a) Feltételezzük, hogy az olaj források megtalálásának valószínűsége **Poisson-eloszlással** modellezhető.

(a1) Rajzoljuk fel az eloszlást és az eloszlásfüggvényt!

(a2) Mekkora a valószínűsége, hogy 1000 fúrásból pontosan 3 lesz sikeres?

(b) Feltételezzük, hogy az olaj források megtalálásának valószínűsége **Binomiális-eloszlással** modellezhető.

(b1) Rajzoljuk fel az eloszlást és az eloszlásfüggvényt!

(b2) Mekkora a valószínűsége, hogy 1000 fúrásból pontosan 3 lesz sikeres?

(c) Hasonlítsuk össze az (a) és (b) részben kapott értékeket!

### 9.10.6. feladat

Egy munkahelyen összesítették 50 munkanapon keresztül, hogy az alkalmazottak naponta hányszor telefonáltak. Egy nap alatti telefonálások száma 0 és 7 között változott az alábbi gyakoriság értékekkel.

Egy nap alatt a telefonálások száma  $\coloneqq [0, 1, 2, 3, 4, 5, 6, 7, 8]$

Napok gyakorisága, ahányszor előfordultak a fenti telefonálási számok  $\coloneqq [2, 7, 12, 11, 9, 5, 3, 1, 0]$

(a) Számítsa ki, hogy naponta átlagosan hányszor telefonálnak az alkalmazottak! ( $\lambda=?$ )

(b) Közelítse a napok gyakoriság értékeit egy **Poisson-eloszlású**  $X$  változóval, ahol  $X$  jelöli az egy nap alatti telefonálások számát!

(c) Közelítse a napok gyakoriság értékeit egy **Binomiális-eloszlású**  $Y$  változóval, ahol  $Y$  jelöli az egy nap alatti telefonálások számát!

(d) Rajzolja fel a mért napok gyakoriságait, a Poisson-modell illetve binomiális modell alapján számolt gyakoriságok oszlop diagrammjaikat közös koordináta-rendszerben! Melyik ad jobb közelítést?



(e) Minimálisan mennyi lesz a napi hívások azon K száma, amelynél több hívás valószínűsége már 0.1-nél kisebb? Számoljon a Poisson-eloszlás eloszlással!

(f) Számolja az (e) feladatot Markov -egyenlőtlenséggel! Összhangban vannak-e az (e) és (f) feladatok megoldásai?

#### 9.10.7. feladat

Legyen X és Y független Bernoulli -féle diszkrét valószínűségi változó, melyekre  $P(X=0) = p_x$  és  $P(Y=0) = p_y$ . Ekkor eloszlásuk

$$\begin{bmatrix} X & 0 & 1 \\ P & p_x & 1 - p_x \end{bmatrix} \quad \begin{bmatrix} Y & 0 & 1 \\ P & p_y & 1 - p_y \end{bmatrix}$$

Határozzuk meg

- (a)  $Z=X+Y$  összegük;
- (b)  $U=X - Y$  különbségük;
- (c)  $V=X \cdot Y$  szorzatuk

eloszlását, várható értékét és varianciáját!

Rajzoljuk fel az eloszlásokat, ha  $p_x = 0.7$  és  $p_y = 0.2$  !

#### 9.10.8. feladat

Egy optikai lemezeket (CD,DVD) gyártó cég  $p=0.01$  valószínűséggel gyárt selejt lemezt. A vállalat 10-s csomagban árulja a lemezeket és csere garanciát ajánl a vásárlóknak, ha 1-nél több lemez hibás a dobozban.

Határozzuk meg annak valószínűségét, hogy egy megvásárolt csomagot ki kell cserélni!

#### 9.10.9. feladat

Egy vállalat azt állítja, hogy annak valószínűsége, hogy egy speciális fedő réteg korrózió álló (egy standard teszt alapján) 0.95. Kiválasztanak egymástól függetlenül 20 elemű mintát.

(a) Ha a fedő réteg olyan jó, mint ahogyan azt állítják, akkor várhatóan hány darab lehet hibás a mintában?

(b) Mekkora a valószínűsége annak, hogy a mintában egynél több hibás lesz?

#### 9.10.10. feladat

Legyen az X valószínűségi változó **Poisson-eloszlású**, melyre  $P(X=0) = 0.0498$  .

Számítsuk ki

- (a) az  $E(X)$  várható értéket és
- (b) a  $P(3 \leq X)$  valószínűséget!

#### 9.10.11. feladat

Egy digitális adatátviteli rendszerben egy bit tévesztésének valószínűsége  $p=10^{-6}=0.000001$  .

Határozzuk meg annak valószínűségét, hogy  $10^6 = 100000$  bitből 3 vagy több bit lesz hibás!

Használjunk a közelítéshez **Poisson -eloszlást**!

#### 9.10.12. feladat

Van n darab villanykörte, melyek között van néhány selejtes. Csak úgy tudjuk megmondani, hogy melyik jó és melyik selejtes, ha becsavarjuk az izzókat egyesével egy foglalatba és felkapcsoljuk az áramot. Tudjuk, hogy egy lámpa  $p$  valószínűséggel selejtes ( $0 < p < 1$ ). Addig vizsgáljuk át a lámpákat egyesével ( és a rosszat eldobjuk), amíg egy jót nem találunk.

Jelölje X az átvizsgált lámpák számát, mire megtaláltuk az első jót!

- (a) Készítsük el X érték- és valószínűség táblázatát  $n=10$  esetén! Rajzoljuk fel az eloszlás pálcika diagramját!
- (b) Számoljuk ki, hogy várhatóan mennyi lámpát kell átvizsgálni tetszőleges n és p esetén!
- (c) Rajzoljuk fel a várható értéket  $n=100$  esetén p függvényében!
- (d) Mekkora p valószínűség mellett lesz X várható értéke 2, ha  $n=100$ ?

#### 9.10.13. feladat

Egy teszt 25 kérdést tartalmaz, amelyek mindegyikére 4 lehetőség közül tudunk választani. A

négy lehetőség közül pontosan egy a helyes. Tegyük fel, hogy egy diák minden kérdés esetén csak találgat.

Jelölje  $X$  a 25 kérdés közül az eltalált helyes válaszok számát!

- (a) Rajzoljuk fel  $X$  eloszlását!
- (b) Mekkora a valószínűsége, hogy a diák 20-nál több kérdésre jól válaszol?
- (c) Mekkora a valószínűsége, hogy a diák 5-nél kevesebb kérdésre válaszol jól?
- (d) Rajzoljuk fel  $X$  eloszlásfüggvényét!
- (e) Várhatóan hány kérdésre válaszol jól a diák?

#### 9.10.14. feladat

Egy bevásárló központban a bank kártyával fizetők közül  $N=1000$  vásárlót választottak ki véletlenszerűen. Azt tapasztalták, hogy 700 olyan vásárló volt, akik egy nagy cég valamelyik termékéből vásároltak az utóbbi 3 hónapon belül. A cég egy új termék bevezetését úgy teszteli, hogy kiválaszt a fenti 1000 vásárló közül  $n=50$  vásárlót véletlenszerűen és minta terméket küld nekik.

(a) Mekkora a valószínűsége, hogy az 50 kiválasztott vásárló közül 45-nél többen vásároltak az utóbbi 3 hónapon belül?

(a1) Számoljunk hipergeometrikus eloszlással!

(a2) Számoljunk binomiális eloszlással!

(b) Hasonlítsuk össze a hipergeometrikus és a binomiális eloszlást az eloszlás diagramok grafikonjai alapján!

#### 9.10.15. feladat

Tegyük fel, hogy egy népszerű rádió műsort telefonon keresztül  $p=0.02$  valószínűséggel tudunk felhívni, vagyis  $q=1-p=0.98$  relatív gyakorisággal jelez a vonal foglalt jelzést. Tegyük fel, hogy az egymást követő hívásaink függetlenek.

(a) Mekkora a valószínűsége, hogy a tizedik hívási kísérletünk lesz az első sikeres hívás?

(b) Mekkora a valószínűsége, hogy 5-nél több hívás szükséges a sikeres kapcsolathoz?

(c) Mennyi hívás szükséges átlagosan az első kapcsolat bekövetkezéséig?

(d) Rajzoljuk fel az első sikeres hívás eloszlását!

#### 9.10.16. feladat

A következő példa tartalmazza Adams által 1936-ban közölt mérést, amely egy útszakaszon 10 másodperces időközönként áthaladó járművek számát és ezek megfigyelt gyakoriságait mutatja.

Hivatkozás: Adams, William F., "Road Traffic Considered as a Random Series," Institution of Civil Engineers Journal, Nov., 1936, pp. 121-13.

| A járművek száma,<br>amelyek egy 10 sec-os<br>periódus alatt érkeztek | A megfigyelt<br>10- sec-os<br>esetek<br>gyakorisága |
|-----------------------------------------------------------------------|-----------------------------------------------------|
| 0                                                                     | 94                                                  |
| 1                                                                     | 63                                                  |
| 2                                                                     | 21                                                  |
| 3                                                                     | 2                                                   |
| >3                                                                    | 0                                                   |

(a) Határozzuk meg a 10 másodpercenként átlagos áthaladó autók számát!

(b) Közelítsük  $X$  **Poisson eloszlással** a 10 másodperc alatt áthaladó autók számát! Rajzoljuk fel a mért és számolt gyakoriságok oszlop diagramjait!

(c) Közelítsük  $Y$  **binomiális eloszlással** a 10 másodperc alatt áthaladó autók számát! Hasonlítsuk össze a mért és számolt gyakoriságok oszlop diagramjait, mind a Poisson, mind a binomiális esetben!

(d) Mennyi a valószínűsége, hogy 1 óra alatt több, mint 240 autó érkezik? Számoljunk Poisson-eloszlással!

#### 9.10.17. feladat

Optikai tároló lemezen (CD, DVD) előforduló szennyeződési hibák **Poisson eloszlást** követnek, átlagosan 1 részecske  $10\text{ cm}^2$ -enként.

Egy lemezen  $100\text{ cm}^2$  területet vizsgálunk át.

(a) Mekkora a valószínűsége, hogy pontosan 12 szennyeződési hiba fordul elő?

(b) Számoljuk ki **binomiális eloszlással** a fenti valószínűséget!

(c) Hasonlítsuk össze a két valószínűségi eloszlás grafikonját!

#### 9.10.18. feladat

A repülőgépek gyakran több helyre adnak el jegyet, mint ahány utas számára van fenntartott ülőhely. Számítanak ugyanis arra, hogy nem minden utas jelenik meg időben az utazáshoz.

Tegyük fel, hogy egy repülőjáratra 125 jegyet adnak el, amelyen csak 120 utas részére van ülőhely. Annak valószínűsége, hogy egy jeggyel rendelkező utas nem tud megjelenni időben legyen  $p=0.10$ , és az utasok viselkedése legyen egymástól független.

(a) Mekkora a valószínűsége annak, hogy az összes megjelent utas fel tud szállni a repülőgépre?

(b) Mekkora a valószínűsége annak, hogy marad üres szék egy járaton?

(c) Átlagosan hány jeggyel rendelkező utas nem száll fel a repülőre.

**Tipp:** Jelöljük  $X$ -szel azoknak az utasoknak a számát, akiknek van ugyan jegye, de nem tudnak megjelenni időben az utazáshoz, valamilyen oknál fogva.

#### 9.10.19. feladat

Egy tétel árú harmadrésze első osztályú. Négy darabot egyesével kiválasztunk a tételből találmra. A kiválasztás és átvizsgálás után mindegyik árút visszatesszük a többi közé, majd úgy húzzuk ki a következőt. Az  $X$  valószínűségi változó értéke legyen a kiválasztott első osztályú darabok száma.

(i) Ábrázoljuk  $X$  eloszlását!

(ii) Számítsuk ki  $X$  várható értékét és szórását!

(iii) Számítsuk ki a  $P(X>2)$  valószínűséget!

#### 9.10.20. feladat

Egy tétel árú 1%-a selejtes. Hány darabot kell találmra kivenni és megvizsgálni, hogy a megvizsgált darabok között legalább 0.95 valószínűséggel selejtes is legyen, ha a kivett darabokat a vizsgálat után visszatesszük.

#### 9.10.21. feladat

Egy automata gépnél megfigyelték, hogy naponta átlagosan 12 db termék lesz selejtes, ezek számának szórása 3.41 darab.

(i) Hány terméket készít naponta a gép?

(ii) Mekkora a valószínűsége, hogy egy napon a selejtes termékek száma 10-nél kevesebb?

#### 9.10.22. feladat

Egy alkatrész halmazból 6 elemű mintát vettünk visszatevéssel. Annak valószínűsége, hogy a minta 3 db selejtet tartalmaz 16%.

Mekkora a selejtarány?

#### 9.10.23. feladat

Egy forgalmas posta hivatalban egy év alatt 1017 címzetlen levelet adtak fel. Mekkora a valószínűsége, hogy egy nap 2-nél több címzés nélküli levelet adnak fel?

#### 9.10.24. feladat

Egy elektronikai műszer 1000 alkatrészből áll. Egy alkatrész a többitől függetlenül 0.001 valószínűséggel romlik el egy év alatt.

Mennyi a valószínűsége, hogy legalább két alkatrész elromlik egy év alatt?

#### 9.10.25. feladat

Egy orsózó gépen 100 munkaóra alatt átlagosan 3 fonal szakadás következik be. Mennyi a

valószínűsége, hogy egy ilyen időtartam alatt a szakadások száma nem lépi túl az átlagot?

#### 9.10.26. feladat

Egy telefon központhoz 600 előfizető tartozik. Tegyük fel, hogy 0.005 a valószínűsége annak, hogy valamelyik előfizető egy meghatározott órában kapcsolást kér. Mennyi a valószínűsége, hogy egy vizsgált óra leforgása alatt éppen 4 előfizető kér vonalat?

#### 9.10.27. feladat

Kalács sütésekor 1 kg tésztába 30 szem mazsolát tesznek egyenletesen elkeverve. Mennyi a valószínűsége, hogy egy 5 dkg -os szeletben 2-nél több mazsola kerül?

#### 9.10.28. feladat

Egy ruhaszövet anyagában 100 méterenként átlagosan 5 hiba keletkezik. Egy 300 méteres szövetet 3 méteres darabokra vágják. Előreláthatólag hány hibátlan darab lesz ezek között?

#### 9.10.29. feladat

Péter bowling játéka során  $p=60\%$ -ban sikerül "tarolást" elérni, vagyis egy golyó gurítással leütni mind a 10 bábút.



(a) Péter  $n=4$  játékot játszik és jelölje  $X$  a tarolások számát a 4 játékból. Rajzolja fel  $X$  eloszlását!

(b) Mekkora a valószínűsége, hogy a 4 játékból legalább 3 esetben ér el tarolást, azaz  $P(3 \leq X) = p_3$ ?

(c) Mekkora a valószínűsége, hogy a 4 játékból nem ér el tarolást, azaz  $P(X=0)=p_0$ ?

(d) Hányszor valószínűbb, hogy legalább 3 esetben ér el tarolást, mint annak hogy nem ér el egyet

sem, azaz  $\frac{p_3}{p_0} = ?$

(e) Legalább mennyi  $n$  játékot kell Péternek játszani, hogy 90%-os valószínűséggel a tarolások száma legalább 3 legyen, azaz  $P(3 \leq Y)=0.9$ , ahol  $Y$  az  $n$  játékból elért tarolások száma.

(f) Az (e) pontban számolt  $n$  érték szemléltetéséhez rajzolja fel az  $(1 - F_Y(3))=P(3 \leq Y)$

függvényt és a 0.9 valószínűségi szintet, az  $n$  játszma függvényében

#### 9.10.30. feladat

Egy repülőtérre utasokat szállító busz üléshelyeinek száma 20 és maximálisan 30 utast szállíthat. A busz menetrendszerint az induló és a végállomás (ami a repülőtér) között csak egy helyen áll meg és kizárólag végállomásig utazó személyeket vesz fel. Megfigyelések alapján az induló állomáson felszálló utasok  $X$  száma Poisson - eloszlást követ  $\mu_1 = 12$  fő utassal. A közbenső megállóban

felszálló utasok  $Y$  száma Poisson - eloszlást követ  $\mu_2 = 10$  fő utassal. A két felszállóban felszálló utasok száma egymástól független.

(a) Rajzoljuk fel  $X$ ,  $Y$  és  $Z = X + Y$  valószínűségi változók eloszlását az első 40 utasra!

(b) Mekkora a valószínűsége, hogy a kezdő megállóban lesz olyan utas, akinek nem jut ülőhely?

(c) Mekkora a valószínűsége, hogy a közbenső megálló után mindenkinek lesz ülőhelye?

(d) Mekkora a valószínűsége, hogy minden utazni szándékozó utast el tudnak szállítani a reptérre?

**Megjegyzés:** Használjuk ki, hogy Poisson-eloszlások összege is Poisson eloszlás!

### 9.10.31. feladat

Egy radioaktív részecskéket számláló Geiger-készüléket úgy helyeztek el egy radioaktív anyag közelében, hogy egy kisugárzott részecskét  $p = \frac{1}{1000} = 0.001$  valószínűséggel detektál vagy

érezkel. Tegyük fel, hogy a vizsgált időtartam alatt  $n=3\,000$  részecskét sugárzott ki a radioaktív anyag.

(a) Mennyi részecskét érezkel a Geiger-készülék legnagyobb valószínűséggel a 3000-ből?

Rajzoljuk fel az érzékelt részecskék számának valószínűségi eloszlását! Használjuk a számolás és rajzolás során az  $X$  binomiális -eloszlást!

(b) Mennyi részecskét érezkel a Geiger-készülék legnagyobb valószínűséggel a 3000-ből?

Rajzoljuk fel az érzékelt részecskék számának valószínűségi eloszlását! Használjunk a számolás és rajzolás során azt az  $Y$  Poisson -eloszlást, melynek várható értéke megegyezik az  $X$  binomiális eloszlás várható értékével!

(c) Hasonlítsuk össze az (a) és (b) pontban kapott eredményeket!

(d) Mekkora a valószínűsége, hogy a Geiger-számláló a 3000 részecskéből **nulla** részecskét érezkel? Számoljunk binomiális és Poisson -eloszlással is!

(e) Mekkora a valószínűsége, hogy a Geiger-számláló a 3000 részecskéből **három** részecskét érezkel? Számoljunk binomiális és Poisson -eloszlással is!

(f) Mekkora a valószínűsége, hogy a Geiger-számláló a 3000 részecskéből **5-nél több** részecskét érezkel? Számoljunk binomiális és Poisson -eloszlással is!

### 9.10.32. feladat

Megfigyelések szerint valamely városban az egy héten előforduló  $X$  balesetek száma Poisson -eloszlást mutat. Statisztikai adatok alapján  $p = 0.6$  annak valószínűsége, hogy valamely héten 1, 2 vagy 3 baleset fordul elő, vagyis  $0.6 = P(X=1)+P(X=2)+P(X=3)$

(a) Ezen feltétel ismeretében, számoljuk ki az 1 héten átlagosan előforduló balesetek  $\lambda$  számát?

Mutassuk meg, hogy a feladatnak két megoldása is van! Számítsuk ki mindkettő  $\lambda_1$  és  $\lambda_2$  paraméter közelítő értékét!

(b) Rajzoljuk fel mindkét Poisson -eloszlást azonos koordinátarendszerben!

(c) Melyik esetben lesz nagyobb annak valószínűsége, hogy egy adott hét alatt nem lesz baleset?

A továbbiakban számoljunk a kisebbik  $\lambda$  paraméterrel!

(d) Mennyi valószínűséggel lesz egy 4 hétből álló hónap alatt 7-nél több baleset?

(e) Hány hetet kell várni arra, hogy a 10-nél több baleset esélye legalább 50% legyen?

Szemléltessük a számítást grafikon felrajzolásával!

### 9.10.33. feladat

Egy víztisztító telepen figyelik a vízben előforduló baktérium telepek számát. A telepen két különálló tartályban végeznek tisztítást. Mindkét tartályban a baktérium telepek száma 1 liter vízben Poisson-eloszlást mutat. Az egyikben literenként  $\lambda_1=5$ , a másikban  $\lambda_2 = 4$  baktérium telep van.

Jelölje  $X_1$  = első tartály 1 literében található baktérium telepek számát és  $X_2$  = második tartály 1 literében található baktérium telepek számát!

(a) Rajzoljuk fel mindkét Poisson -eloszlást azonos koordinátarendszerben!

(b) Melyik esetben lesz nagyobb annak valószínűsége, hogy a telepek száma egy literben nem lesz több 1-nél?

(c) Összeöntünk egyforma mennyiségű vizet a két tartályból. Milyen lesz az összeöntött és elkevert víz egy literjében található baktérium telepek számának eloszlása? Mekkora

valószínűséggel lesz az összeöntött víz egy literjében 1 vagy 1-nél kevesebb baktérium telep?

(d) Milyen  $0 < t < 1$  arányban kell az első tartályból és  $(1-t)$  arányban a második tartályból a vizet összeönteni, hogy az összeöntött és elkevert vízben levő baktérium telepek száma  $p=0.05$  valószínűséggel ne legyen több, mint 1? Szemléltessük a számítást a  $0 < t < 1$  keverési arány

grafikonjának és a  $p=0.05$  valószínűségi szint felrajzolásával!

#### 9.10.34. feladat

A világ kereskedelmének 92%-a tengeren zajlik. Közel 80 000 olyan kereskedelmi hajó járja a világ minden részét, amelyek átlagosan  $\lambda=500$  tonna súlyt szállítanak. Minden évben több, mint 120 hajó tűnik el ezek közül a tengerben.

Tegyük fel, hogy egy évben pontosan  $N=80\,000$  kereskedelmi hajó indul útnak és ezek közül  $n=120$  tűnik el az adott évben.

(a) Egy nagy hajózási cég  $F = 160$  kereskedelmi hajóból álló flottát működtet. Jelölje  $T$  azon hajók számát a cég hajói közül, amelyek az adott évben eltűnnek.

Milyen a  $T$  véletlen változó eloszlása? Mekkora valószínűséggel tűnik el  $T=2$  hajó egy évben? Milyen feltételezéssel éltünk a számítás során?

(b) Rajzoljuk fel a hajók eltűnési valószínűségeit pálcika diagrammal  $T=0, 1, 2, 3$  eltűnési számokra!

Legnagyobb valószínűséggel hány hajó tűnik el a flottából egy év alatt?

(c) Várhatóan hány hajója tűnik el évente a cégnek? Hány év alatt tűnik el 5 hajó?

(d) Egy hajó által szállított  $S$  súly átlagosan  $\mu=500$  tonna, melynek eloszlása normál  $\sigma = 30$  tonna szórással. Rajzoljuk fel  $S$  sűrűség függvényét!

(e) Milyen eloszlású a flotta  $F=160$  független hajója által összesen szállított  $Z$  teljes súly?

Rajzoljuk fel  $Z$  sűrűség függvényét! (Legyenek az egyes hajók szállítmányainak súlya egymástól független!)

#### 9.10.35. feladat

Egy halászatot folytató tengeri flotta a tonhal telepek felkutatására radar kereső berendezést használ. Feltételezhető, hogy a tonhal telepek az adott területen véletlenszerűen helyezkednek el, melynek átlagos előfordulási sűrűsége  $A= 200\,000\text{ km}^2$ -enként 1 telep. A radar naponta  $T=24\,000\text{ km}^2$  területet tud átvizsgálni.

Mekkora a valószínűsége, hogy a radar talál legalább egy tonhal telepet 5 napos keresés során?

(a) Használjunk  $X$  Poisson eloszlást az 5 nap alatt található tonhal telepek számára! A számítást végezzük az  $X$  változó eloszlásával!

(b) Használjunk  $Y$  Binomiális eloszlást az 5 nap alatt található tonhal telepek számára! A számítást végezzük az  $Y$  változó eloszlásával!

(c) Számítsuk az (a) és (b) pontokban kapott valószínűségek eltéréseinek relatív hibáját!

A flotta menedzsmenete a költségek (üzemanyag stb..) várható alakulása miatt szeretné tudni, hogy hány nap szükséges ahhoz, hogy legalább egy tonhal telepre bukkanjanak.

(d) Poisson eloszlással számolva keressük meg azon napok  $n$  számát, amely elegendő ahhoz, hogy  $p=90\%$ -os biztonsággal találjanak legalább egy tonhal telepet! A számítást végezzük a Poisson-eloszlású  $Z_n$  változóval, mely az " $n$ " nap alatt található tonhal telepek számát adja meg.

(e) Rajzoljuk fel a  $Z_n$  változóra a  $P(1 \leq Z_n)$  valószínűséget az  $n$  nap függvényében és a  $p=0.9$  valószínűségi szintet!

#### 9.10.36. feladat

Egy nagy városba vezető forgalmas úton átlagosan 10 percenként reggelente 3 autó megy befelé a városba és átlagosan 2 autó hagyja el a várost 10 percenként.

A 10 perc alatt beérkező autók  $X$  száma és a várost elhagyó autók  $Y$  száma modellezhető független Poisson-eloszlásokkal.

(a) Rajzoljuk fel az  $X$  beérkező és  $Y$  távozó autók  $P(X = n)$  és  $P(Y=n)$  valószínűségi eloszlásait  $n=0,1,2,3,\dots,10$  értékekre!

(b) Melyik valószínűség nagyobb: 10 perc alatt nem érkezik befelé egy autó sem, vagy 10 perc alatt nem hagyja el a várost egy autó sem: azaz  $P(X=0)$  vagy  $P(Y=0)$ ?

(c) Határozzuk meg annak valószínűségét, hogy valamely 10 perc alatt egyik irányba sincs mozgás!

(d) Mekkora annak valószínűsége, hogy 10 perc alatt összesen maximum 3 autó halad át az úton?



(e) Rajzoljuk fel az  $X$  és  $Y$  autókra a  $P(X \leq n)$  és  $P(Y \leq n)$  eloszlás függvényeket  $n=0,1,2,3,\dots,10$  értékekre, amelyek a 10 perc alatt legfeljebb " $n$ " áthaladás valószínűségei az egyes irányokba! Melyik lépcsős függvény értékei nagyobbak!

(f) Mekkora a valószínűsége, hogy egy 1 óra alatt 40-nél több autó halad el összesen az úton?

#### 9.10.37. feladat

Az elveszett tárgyak osztályán őrzik egy főpályaudvarra beérkező vonaton hagyott tárgyakat, két kategóriába sorolva..

Az "A" kategóriába tartoznak a ruhaneműk (kabát, kesztyű, esernyő, stb..), "B" kategóriába tartoznak a papír árúk (dosszié, könyv, papír, stb...).

Egy bizonyos időtartam alatt átlagosan naponta 2 "A" kategóriába tartozó tárgyat hagytak el, míg naponta átlagosan 3 "B" kategóriába tartozó tárgyat hagytak el.

Az elhagyott tárgyak napi száma véletlen változó, melyek modellezhetők **Poisson-eloszlással** mindkét kategóriában. Két különböző napon elvesztett tárgyak száma egymástól független.

(a) Rajzoljuk fel az "A" és "B" kategóriákban a  $P(A = n)$  és  $P(B = n)$  valószínűségi eloszlásokat  $n=0,1,2,3,\dots,10$  értékekre, melyek az egy nap alatt elhagyott " $n$ " tárgy valószínűségei az egyes kategóriában! (2 p.)

(b) Melyik valószínűség nagyobb: egy nap alatt nem hagynak el "A" kategóriás tárgyat vagy egy nap alatt nem hagynak "B" kategóriás tárgyat a vonatokon: azaz  $P(A=0)$  vagy  $P(B=0)$ ? (2 p.)

(c) Határozzuk meg annak valószínűségét, hogy egy nap alatt nem hagynak el semmit a vonatokon! (2 p.)

(d) Mekkora annak valószínűsége, hogy egy nap alatt összesen 3 elhagyott tárgyat találnak a vonatokon? (2 p.)

(e) Rajzoljuk fel az "A" és "B" kategóriákban a  $P(A \leq n)$  és  $P(B \leq n)$  eloszlás függvényeket  $n=0,1,2,3,\dots,10$  értékekre, amelyek az egy nap alatt legfeljebb " $n$ " elhagyott tárgy valószínűségei az egyes kategóriákban! Melyik lépcsős függvény értékei nagyobbak! (2 p.)

(f) Mekkora a valószínűsége, hogy egy hét alatt 16-nál több "A" kategóriás tárgyat hagynak el?

#### 9.10.38. feladat

Egy ruházati boltban trikókat árulnak, háromféle méretben: S-, M- és L méretűt és kétféle

színben fehér és kék. Tegyük fel, hogy az S méretűek  $p[S] = \frac{1}{4}$  része, az M méretűek,  $p[M]$

$= \frac{3}{5}$  része és az L méretűek  $p[L] = \frac{1}{3}$  része fehér színű! Kiválasztunk 3 sokaságokból

véletlenszerűen egy S, egy M és egy L méretű trikót, egymástól függetlenül.

Jelölje  $X$  a kék színű trikók számát a három közül.

(a) Adjuk meg  $X$  lehetséges értékeit és a hozzá tartozó valószínűségeket!

(b) Ábrázoljuk  $X$  eloszlását! Mennyi kék trikó lesz a 3 között legnagyobb valószínűséggel

(c) Számítsuk ki  $X$  várható értékét és szórását!

#### 9.10.39. feladat

Megfigyelések szerint valamely városban az egy héten előforduló  $X$  balesetek száma Poisson - eloszlást mutat. Statisztikai adatok alapján  $p = 0.6$  annak valószínűsége, hogy valamely héten 1, 2 vagy 3 baleset fordul elő, vagyis  $0.6 = P(X=1)+P(X=2)+P(X=3)$

(a) Ezen feltétel ismeretében, számoljuk ki az 1 héten átlagosan előforduló balesetek  $\lambda$  számát?

Mutassuk meg, hogy a feladatnak két megoldása is van! Számítsuk ki mindkettő  $\lambda_1$  és  $\lambda_2$

paraméter közelítő értékét!

(b) Rajzoljuk fel mindkét Poisson -eloszlást azonos koordináta-rendszerben!

(c) Melyik esetben lesz nagyobb annak valószínűsége, hogy egy adott hét alatt nem lesz baleset? A továbbiakban számoljunk a kisebbik  $\lambda$  paraméterrel!

(d) Mennyi valószínűséggel lesz egy 4 hétből álló hónap alatt 7-nél több baleset?

(e) Hány hetet kell várni arra, hogy a 10-nél több baleset esélye legalább 50% legyen?

Szemléltessük a számítást grafikon felrajzolásával!

#### 9.10.40. feladat

Egy elektronikai cikket forgalmazó cég úgy gondolja, hogy termékeinek 2%-a romlik el a garancia időn belül. Egy (függetlenül kiválasztott) 500 elemű eladott minta garancia idő alatti életútját követték figyelemmel.

**(a)** Mekkora a valószínűsége, hogy a garancia időszak alatt az  $n = 500$  vevő közül senki sem reklamál?

**(b)** Mekkora a garancia időn belül meghibásodott elektronikai cikkek várható száma?

**(c)** Mekkora a valószínűsége, hogy több, mint 2 elektronikai cikk hibásodik meg a garancia időszakon belül?

**(d)** Rajzoljuk fel az ugyanilyen várható értékű Poisson -eloszlás és a binomiális eloszlás eloszlás grafikonját, közös koordináta-rendszerben!