Cvičení 9: Pravděpodobnostní funkce, hustoty a distribuční funkce v systému STATISTICA, výpočet pravděpodobností pomocí distribučních funkcí

Systém STATISTICA vytváří grafy hustot a distribučních funkcí mnoha spojitých rozložení, umí stanovit hodnotu distribuční funkce či počítat 1 - hodnota distribuční funkce. Slouží k tomu Pravděpodobnostní kalkulátor v menu Statistiky. Hodnoty pravděpodobnostních funkcí, hustot a distribučních funkcí lze počítat též pomocí funkcí implementovaných v položce „Dlouhé jméno" proměnné.
Zaměříme se na binomické rozložení, Poissonovo rozložení, exponenciální rozložení a normální rozložení.

Binomické rozložení Bi(n, ϑ)

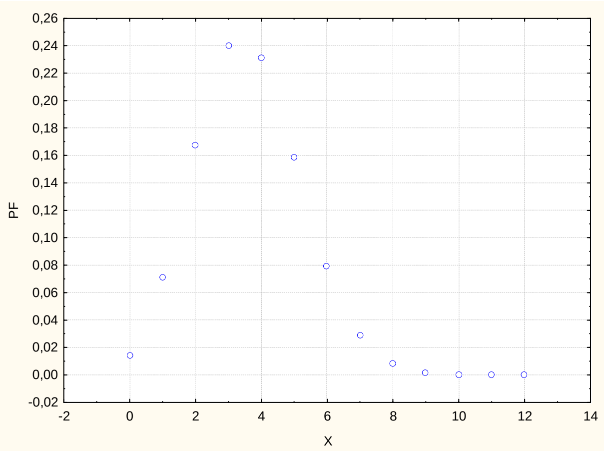
Náhodná veličina X udává počet úspěchů v posloupnosti n nezávislých opakovaných pokusů, přičemž pravděpodobnost úspěchu je v každém pokusu ϑ. Píšeme X ~ Bi(n, ϑ).

$$\pi(x) = \begin{cases} \binom{n}{x} \vartheta^x (1-\vartheta)^{n-x} & \text{pro } x = 0, \dots, n \\ 0 \text{ jinak} \end{cases}, \Phi(x) = \sum_{t=0}^{x} \binom{n}{t} \vartheta^t (1-\vartheta)^{n-t}$$

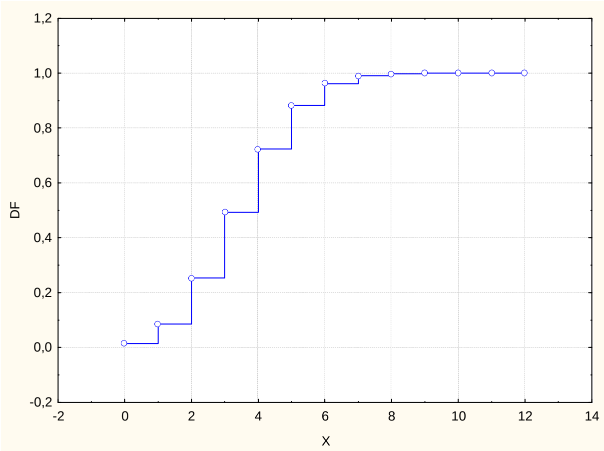
Kreslení grafů funkcí $\pi(x)$ a $\Phi(x)$ v systému STATISTICA

1. možnost: Ukážeme si, jak získat grafy pravděpodobnostní a distribuční funkce náhodné veličiny $X \sim Bi(12;0,3)$. Vytvoříme nový datový soubor o 3 proměnných a 13 případech. První proměnnou nazveme X a uložíme do ní hodnoty 0, 1, ..., 12 (do Dlouhého jména napíšeme =v0-1). Druhou proměnnou nazveme PF a uložíme do ní hodnoty pravděpodobnostní funkce (do Dlouhého jména napíšeme příkaz =Binom(x;0,3;12)). Třetí proměnnou nazveme DF a uložíme do ní hodnoty distribuční funkce (do Dlouhého jména napíšeme příkaz =IBinom(x;0,3;12)).
Graf pravděpodobnostní funkce: Grafy – Bodové grafy – Proměnné X, PF – OK – vypneme Lineární proložení – OK.
Graf distribuční funkce: Grafy – Bodové grafy – Proměnné X, DF – OK – vypneme Lineární proložení – OK – 2x klikneme na pozadí grafu – Graf:Obecné – zaškrtneme Spojnice – Typ spojnice: Schod – OK. (Svislé čáry do grafu nepatří.)

Graf funkce $\pi(x)$ rozložení $Bi(12;0,3)$

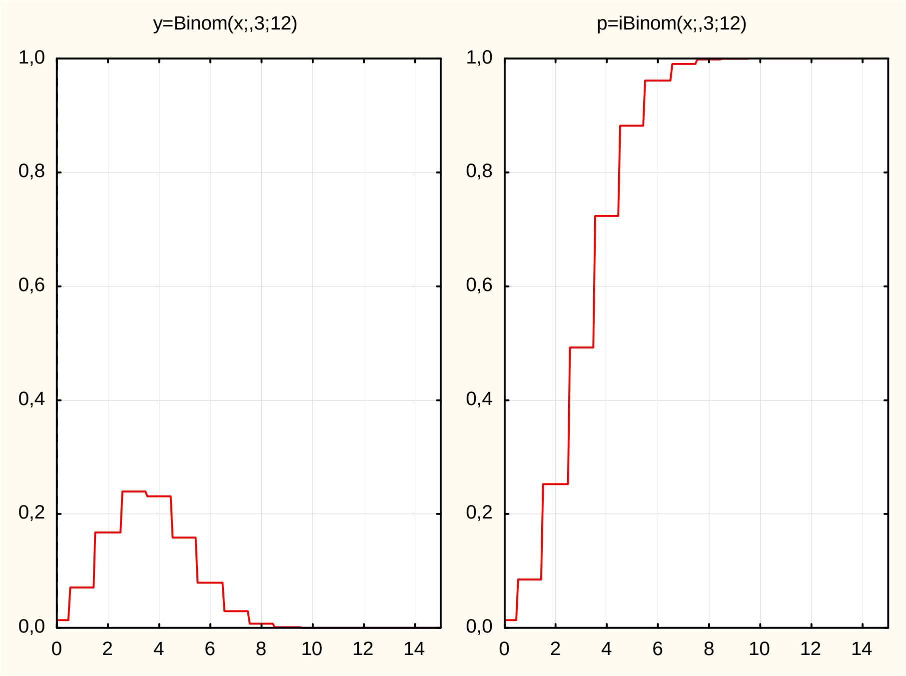


Graf funkce $\Phi(x)$ rozložení $Bi(12;0,3)$



Analogickým způsobem můžeme získat grafy pravděpodobnostních distribučních funkcí binomického rozložení pro různá n a ϑ a sledovat vliv těchto parametrů na vzhled grafů.

2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Binomické. Vyplníme X: 0, N: 12, p: 0,3, zaškrtneme Vytv. graf – Výpočet.



Graf pravděpodobnostní funkce není z formálního hlediska správný, protože pravděpodobnostní funkce je kladná pouze v bodech 0, 1, …, n (=12) všude jinde je nulová. Do grafu distribuční funkce nepatří svislé čáry.

Poissonovo rozložení Po(λ)

Náhodná veličina X udává počet událostí, které nastanou v jednotkovém časovém intervalu (resp. v jednotkové oblasti), přičemž k událostem dochází náhodně, jednotlivě a vzájemně nezávisle. Parametr $\lambda > 0$ je střední počet těchto událostí. Píšeme $X \sim \text{Po}(\lambda)$.

$$\pi(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{pro } x = 0, 1, \dots \\ 0 \text{ jinak} \end{cases}, \Phi(x) = \sum_{t=0}^{x} \frac{\lambda^t}{t!} e^{-\lambda}$$

Kreslení grafů funkcí $\pi(x)$ a $\Phi(x)$ v systému STATISTICA

1. možnost: Při tvorbě grafů pravděpodobnostní a distribuční funkce náhodné veličiny s Poissonovým rozložením, např. $X \sim \text{Po}(5)$, postupujeme podobně jako u binomického rozložení, ale v datovém souboru bude 16 případů a použijeme funkce Poisson(x;5) a IPoisson(x;5).
2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Poisson. Vyplníme X: 0, Lambda 5, zaškrtneme Vytv. graf – Výpočet.

Příklad 1.: Při provozu balicího automatu vznikají během směny náhodné poruchy, které se řídí rozložením Po(2). Jaká je pravděpodobnost, že během směny dojde k aspoň jedné poruše?

Řešení: X – počet poruch během směny, X ~ Po(2), P(X ≥ 1) = 1 – P(X < 1) = 1 – P(X = 0) =
$= 1 - \frac{2^0}{0!}e^{-2} = 0{,}8647$.

Návod na výpočet pomocí systému STATISTICA:

1. možnost: Otevřeme nový datový soubor o jedné proměnné a jednom případu. Do Dlouhého jména této proměnné napíšeme =1-IPoisson(0;2). Dostaneme výsledek 0,8647.

2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Poisson. Vyplníme X: 0, Lambda 2, zaškrtneme Kum. Pravděpodobnost – Výpočet. Dostaneme 0,1353. Hledaná pravděpodobnost je doplněk do 1, tedy 0,8647.

Exponenciální rozložení Ex(λ)

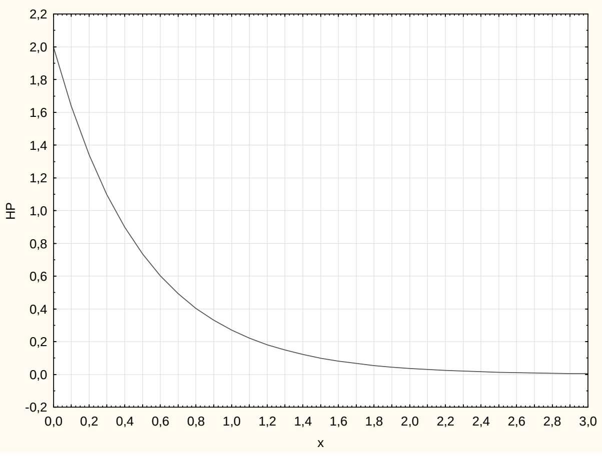
Náhodná veličina X udává dobu čekání na příchod nějaké události, která se může dostavit každým okamžikem se stejnou šancí bez ohledu na dosud pročekanou dobu. (Jde o tzv. čekání bez paměti.) Přitom $\frac{1}{\lambda}$ vyjadřuje střední dobu čekání. Náhodná veličina X ~ Ex(λ) má hustotu

$$\varphi(x) = \begin{cases} \lambda e^{-\lambda x} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases}.$$

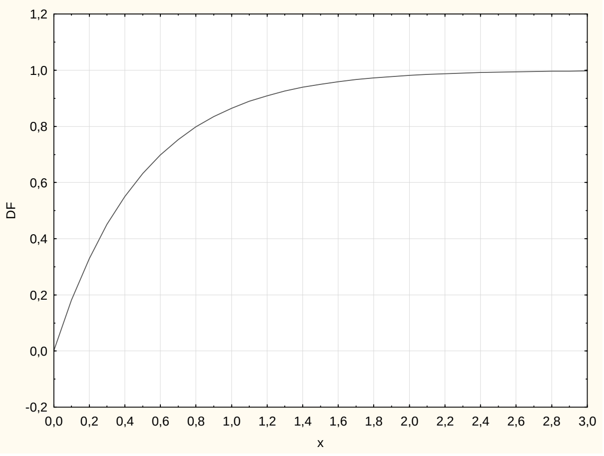
Kreslení grafů funkcí $\varphi(x)$ a $\Phi(x)$ exponenciálního rozložení v systému STATISTICA

1. možnost: Ukážeme si, jak získat grafy hustoty pravděpodobnosti a distribuční funkce náhodné veličiny $X \sim Ex(2)$. Vytvoříme nový datový soubor o 3 proměnných a 30 případech. První proměnnou nazveme X a uložíme do ní hodnoty 0,001,1, 0,2, ..., 3 (do Dlouhého jména napíšeme =v0/10). Před 1. případ vložíme jeden případ a do proměnné X napíšeme 0,001. Druhou proměnnou nazveme HP a uložíme do ní hodnoty hustoty pravděpodobnosti. (do Dlouhého jména napíšeme příkaz =Expon(x;2)). Třetí proměnnou nazveme DF a uložíme do ní hodnoty distribuční funkce (do Dlouhého jména napíšeme příkaz =IExpon(x;2)).
Graf hustoty: Grafy – Bodové grafy – Proměnné X, HP – OK – vypneme Lineární proložení – OK. Vymažeme značky, zapneme spojnici a na vodorovné ose upravíme měřítko od 0 do 3.
Graf distribuční funkce: Grafy – Bodové grafy – Proměnné X, DF – OK – vypneme Lineární proložení – OK. Vymažeme značky, zapneme spojnici a na vodorovné ose upravíme měřítko od 0 do 3.

Graf funkce $\varphi(x)$ rozložení $Ex(2)$



Graf funkce $\Phi(x)$ rozložení $Ex(2)$



2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Exponenciální. Vyplníme lambda: 2, zaškrtneme Vytv. graf – Výpočet.

Příklad 2.: Doba do ukončení opravy v opravně obuvi je náhodná veličina, která se řídí exponenciálním rozložením se střední dobou opravy 3 dny. Jaká je pravděpodobnost, že oprava bude ukončena do dvou dnů?

Řešení: $X \sim Ex(1/3),\; P(X \le 2) = \int_0^2 \frac{1}{3} e^{-\frac{x}{3}} dx = \left[-e^{-\frac{x}{3}} \right]_0^2 = 1 - e^{-\frac{2}{3}} = 0{,}4866$

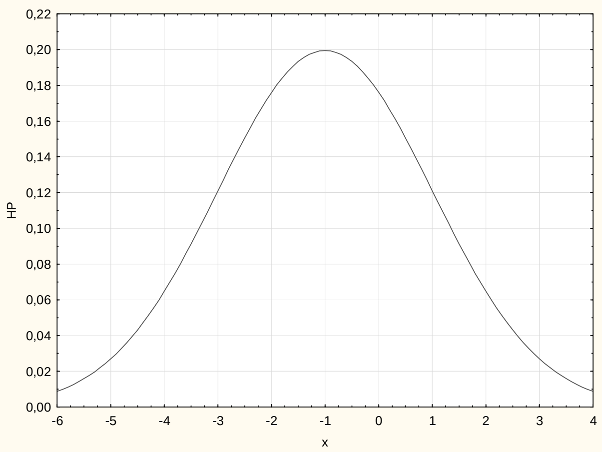
Návod na výpočet pomocí systému STATISTICA:
1. možnost: Otevřeme nový datový soubor o jedné proměnné a jednom případu.
Do dlouhého jména této proměnné napíšeme =IExpon(2;1/3). Dostaneme 0,4866.
2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Exponenciální - do okénka lambda napíšeme 0,3333, do okénka exp. napíšeme 2 a po kliknutí na Výpočet se v okénku p objeví 0,4866.

Normální rozložení $N(\mu, \sigma^2)$

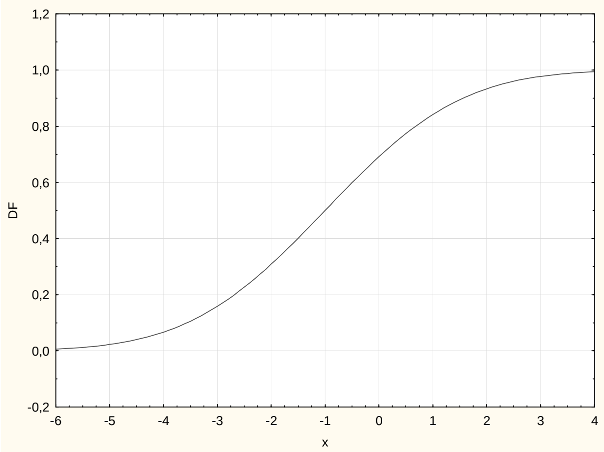
Náhodná veličina $X \sim N(\mu, \sigma^2)$ má hustotu $\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. Pro $\mu = 0$, $\sigma^2 = 1$ se jedná o standardizované normální rozložení, píšeme $U \sim N(0, 1)$. Hustota pravděpodobnosti má v tomto případě tvar $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$.

Kreslení grafů funkcí $\varphi(x)$ a $\Phi(x)$ normálního rozložení v systému STATISTICA
1. možnost: Ukážeme si, jak získat grafy hustoty pravděpodobnosti a distribuční funkce náhodné veličiny $X \sim N(-1{,}4)$. Vytvoříme nový datový soubor o 3 proměnných a 101 případech. První proměnnou nazveme x a uložíme do ní hodnoty -6, -5,9, -5,8, ..., 4 (do Dlouhého jména napíšeme =(v0-1)/10-6). Druhou proměnnou nazveme HP a uložíme do ní hodnoty hustoty pravděpodobnosti. (do Dlouhého jména napíšeme příkaz =Normal(x;-1;2)). Třetí proměnnou nazveme DF a uložíme do ní hodnoty distribuční funkce (do Dlouhého jména napíšeme příkaz = INormal(x;-1;2)).
Graf hustoty: Grafy – Bodové grafy – Proměnné x, HP – OK – vypneme Lineární proložení – OK. Vymažeme značky, zapneme spojnici a na vodorovné ose upravíme měřítko od -6 do 4.
Graf distribuční funkce: Grafy – Bodové grafy – Proměnné x, DF – OK – vypneme Lineární proložení – OK. Vymažeme značky, zapneme spojnici a na vodorovné ose upravíme měřítko od -6 do 4.

Graf funkce $\varphi(x)$ rozložení $N(-1,4)$



Graf funkce $\Phi(x)$ rozložení $N(-1,4)$



2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Z (normální). Vyplníme průměr: -1, SmOdch: 2, zaškrtneme Vytv. graf – Výpočet.

Příklad 3: Životnost baterie v hodinách je náhodná veličina, která má normální rozložení se střední hodnotou 300 hodin a směrodatnou odchylkou 35 hodin. Jaká je pravděpodobnost, že náhodně vybraná baterie bude mít životnost aspoň 320 hodin?

Návod na výpočet pomocí systému STATISTICA:
1. možnost: Otevřeme nový datový soubor o jedné proměnné a jednom případu.
Do dlouhého jména této proměnné napíšeme =1-INormal(320;300;35). Dostaneme 0,2839.
2. možnost: Využijeme Pravděpodobnostní kalkulátor.
Statistiky – Pravděpodobnostní kalkulátor – Rozdělení – Z(normální) - do okénka průměr napíšeme 300, do okénka SmOdch napíšeme 35, zaškrtneme 1-kumul. p a po kliknutí na Výpočet se v okénku p objeví 0,2839.

Příklad k samostatnému řešení.: Na výrobní lince jsou automaticky baleny balíčky rýže o deklarované hmotnosti 1000 g. Působením náhodných vlivů hmotnost balíčků kolísá. Lze ji považovat za náhodnou veličinu, která se řídí normálním rozložením se střední hodnotou 996 g a směrodatnou odchylkou 18 g. Jaká je pravděpodobnost, že náhodně vybraný balíček rýže neprojde výstupní kontrolou, jestliže je povolená tolerance ± 30 g od deklarované hmotnosti 1000 g?
Výsledek:
$$P\left(X \notin \langle 970,1030 \rangle\right) = 1 - P(970 < X < 1030) = 0,104$$