Základní rozdělení pravděpodobnosti

Diskrétní rozdělení pravděpodobnosti

1. Pojem Náhodná veličina s **Binomickým rozdělením** $\mathrm{Bi}(n, p)$, kde n je přirozené číslo, p je reálné číslo, $0 < p < 1$ má pravděpodobnostní funkci

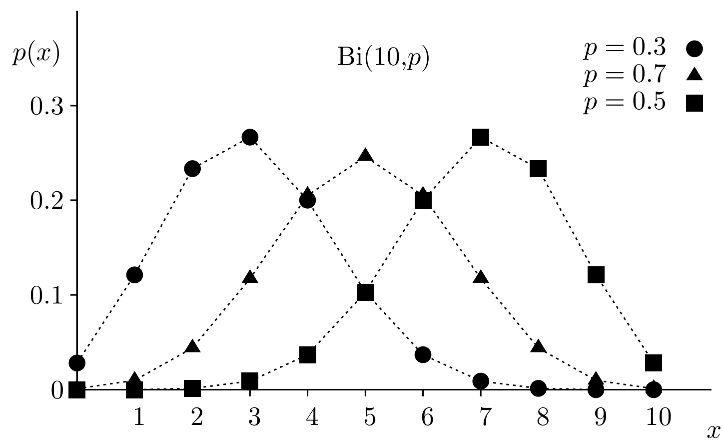
$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$$

2. Vlastnosti Číselné charakteristiky binomického rozdělení jsou

$$\begin{aligned} \mathrm{E}(X) &= np \\ \mathrm{D}(X) &= np(1-p) \\ \mathrm{A}_3(X) &= \frac{1-2p}{\sqrt{np(1-p)}} \\ (n+1)p - 1 \leq \hat{x} &\leq (n+1)p. \end{aligned}$$

Toto rozdělení má náhodná veličina X udávající počet nastoupení sledovaného náhodného jevu v posloupnosti n vzájemně nezávislých pokusů. Jedná se také o popis tzv. ***náhodného výběru s vracením***, kdy např. postupně vybíráme z dodávky n výrobků ke kontrole, X je počet zmetků mezi nimi, p je pravděpodobnost výroby zmetku, a každý vybraný výrobek vracíme zpět do dodávky (to znamená, že může být znovu kontrolován).

3. Poznámka Na Obrázku 1 jsou uvedeny příklady pravděpodobnostních funkcí binomického rozdělení. Pro $p = 0.5$ je binomické rozdělení symetrické a pro $p > 0.5$, resp. $p < 0.5$, je záporně, resp. kladně, asymetrické.



Obrázek 1: Grafy pravděpodobnostní funkce binomického rozdělení pro $n = 10$ a různé hodnoty parametru p. Přerušovaná čára je použita pouze pro odlišení jednotlivých pravděpodobnostních funkcí.

4. Pojem Rozdělení $\mathrm{Bi}(1, p)$, tedy pro $n = 1$, se nazývá **alternativní rozdělení** a značí se $\mathrm{A}(p)$.

5. Vlastnosti Náhodná veličina $X = X_1 + \cdots + X_k$, kde náhodné veličiny X_j, $j = 1, \ldots, k$ jsou nezávislé a mají binomická rozdělení $\mathrm{Bi}(n_j, p)$ se stejným parametrem p, má binomické rozdělení $\mathrm{Bi}(n, p)$, kde $n = n_1 + \ldots + n_k$.

Speciálně součet n nezávislých náhodných veličin s alternativním rozdělením $\mathrm{A}(p)$ má binomické rozdělení $\mathrm{Bi}(n, p)$.

6. Příklad V dodávce 50 výrobků je 5 zmetků. Z dodávky jsou náhodně vybrány 3 výrobky. Počet zmetků mezi vybranými výrobky je náhodná veličina X. Určete typ jejího rozdělení pravděpodobnosti, její pravděpodobnostní funkci $p(x)$, střední hodnotu $\mathrm{E}(X)$, rozptyl $\mathrm{D}(X)$, směrodatnou odchylku $\sigma(X)$, koeficient šikmosti $\mathrm{A}_3(X)$, medián $x_{0.5}$, modus $\hat{x}$ a $P(1 < X \leq 3)$. Předpokládejte, že každý vybraný výrobek se vrátí nazpět do dodávky, takže jde o náhodný výběr s vracením.

Řešení Náhodná veličina X má rozdělení $\mathrm{Bi}(n,p)$, kde $n = 3$ a $p = 5/50 = 0.1$. Náhodná veličina X nabývá hodnot $x = 0, 1, 2, 3$ a její pravděpodobnostní funkce je

$$p(x) = \binom{3}{x} 0.1^x \cdot 0.9^{3-x} \quad \text{pro} \quad x = 0, 1, 2, 3.$$

Ze vzorců můžeme vypočíst následující číselné charakteristiky.

Střední hodnota	$\mathrm{E}(X) = np = 3 \cdot 0.1 = 0.3$,
Rozptyl	$\mathrm{D}(X) = np(1-p) = 3 \cdot 0.1 \cdot 0.9 = 0.27$,
Směrodatná odchylka	$\sigma(X) = \sqrt{\mathrm{D}(X)} = \sqrt{0.27} \approx 0.51962$,
Koeficient šikmosti	$\mathrm{A}_3(X) = \frac{1-2p}{\sqrt{np(1-p)}} = \frac{1-2\cdot 0.1}{\sqrt{0.27}} \approx 2.7245$,
Medián	$x_{0.5} = 0$, neboť $F(x) = p(0) = \binom{3}{0} 0.1^0 \cdot 0.9^{3-0} = 0.729 > 0.5$ pro $x \in (0;1\rangle$,
Modus	$\hat{x} = 0$, neboť $(n+1)p - 1 = -0.6$ a $(n+1)p = 0.4$,

Z pravděpodobnostní funkce pak lze přímo vypočíst pravděpodobnost.

$$P(1 < X \leq 3) = p(2) + p(3) = 0.027 + 0.001 = 0.028.$$

7. Pojem Náhodná veličina s **Hypergeometrickým rozdělením** $\mathrm{H}(N, M, n)$, kde N, M a n jsou přirozená čísla, $1 \leq n \leq N$, $1 \leq M \leq N$ má pravděpodobnostní funkci

$$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \quad x = \max\{0, M - N + n\}, \ldots, \min\{M, N\}.$$

8. Vlastnosti Náhodná veličina s Hypergeometrickým rozdělením má následující číselné charakteristiky

$$\begin{aligned} \mathrm{E}(X) &= n\frac{M}{N}, \\ \mathrm{D}(X) &= n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}, \\ a - 1 \leq \quad \hat{x} &\leq a, \quad \text{kde} \quad a = \frac{(M+1)(n+1)}{N+2} \end{aligned}$$

Hypergeometrické rozdělení popisuje tzv. **náhodný výběr bez vracení**, kdy např. N je celkový počet výrobků, M je počet zmetků mezi těmito výrobky a vybereme náhodně (bez vracení jednotlivých výrobků nebo jejich skupin) celkem n výrobků, mezi nimiž je x zmetků.

9. Poznámka Na Obrázku 2 jsou uvedeny příklady pravděpodobnostních funkcí hypergeometrického rozdělení.

10. Příklad V dodávce 50 výrobků je 5 zmetků. Z dodávky jsou náhodně vybrány 3 výrobky. Počet zmetků mezi vybranými výrobky je náhodná veličina X. Určete typ jejího rozdělení pravděpodobnosti, její pravděpodobnostní funkci $p(x)$, střední hodnotu $\mathrm{E}(X)$, rozptyl $\mathrm{D}(X)$, směrodatnou odchylku $\sigma(X)$, medián $x_{0,5}$, modus $\hat{x}$ a $\mathrm{P}(1 < X \leq 3)$. Předpokládejte (na rozdíl od řešeného příkladu 6), že se vybraný výrobek nevrací nazpět do dodávky, takže jde o náhodný výběr bez vracení.

Řešení Náhodná veličina X má rozdělení $\mathrm{H}(N, M, n)$, kde $N = 50$, $M = 5$ a $n = 3$. Náhodná veličina

X nabývá hodnot $x = 0$, 1, 2, 3 a její pravděpodobnostní funkce je

$$p(x) = \frac{\binom{5}{x}\binom{45}{3-x}}{\binom{50}{3}} \quad \text{pro} \quad x = 0, 1, 2, 3.$$

Ze vzorců můžeme vypočíst následující číselné charakteristiky.

Střední hodnota	$\mathrm{E}(X) = n\frac{M}{N} = 3 \cdot 0.1 = 0.3$,
Rozptyl	$\mathrm{D}(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1} = 3 \cdot 0.1 \cdot 0.9 \cdot (47/49) \doteq 0.25898$,
Směrodatná odchylka	$\sigma(X) = \sqrt{\mathrm{D}(X)} \doteq \sqrt{0.25898} \doteq 0.50890$,
Medián	$x_{0.5} = 0$, neboť $F(x) = p(0) = \frac{\binom{5}{0}\binom{45}{3-0}}{\binom{50}{3}} \doteq 0.72398 > 0.5$, pro $x \in (0, 1\rangle$
Modus	$\hat{x} = 0$, neboť $a = \frac{(M+1)(n+1)}{N+2} \doteq 0.46154$, takže $a - 1 \doteq -0.53846$,

Z pravděpodobnostní funkce pak lze přímo vypočíst pravděpodobnost

$$P(1 < X \leq 3) = p(2) + p(3) \doteq 0.02296 + 0.00051 = 0.02347.$$

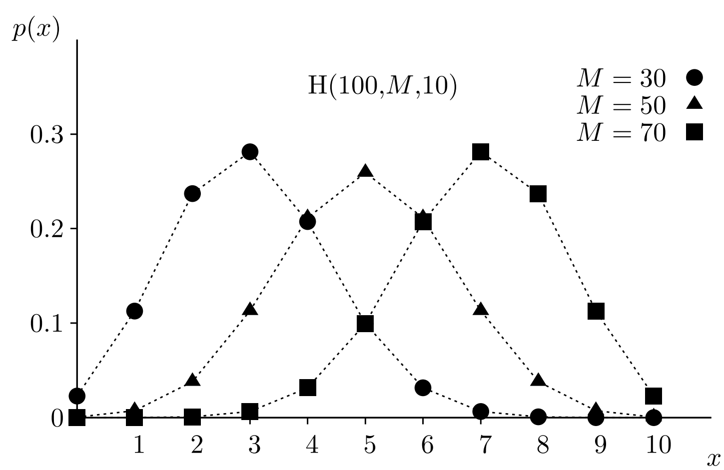
11. Pojem Náhodná veličina s **Poissonovým rozdělením** Po(λ), kde λ je reálné číslo, $\lambda > 0$, má pravděpodobnostní funkci

$$p(x) = \frac{\lambda^x}{x!}e^{-\lambda}, \quad x = 0, 1, \ldots$$

12. Vlastnosti Náhodná veličina s Poissonovým rozdělením má následující číselné charakteristiky

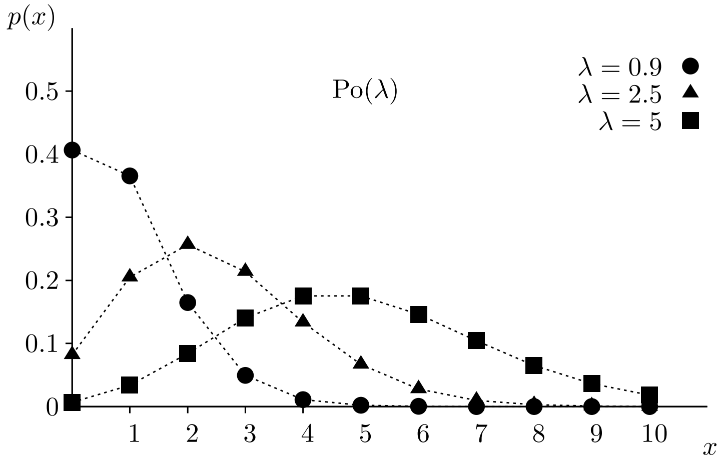
$$\begin{aligned} \mathrm{E}(X) &= \lambda, \\ \mathrm{D}(X) &= \lambda, \\ \mathrm{A}_3(X) &= \frac{1}{\sqrt{\lambda}} \\ \lambda - 1 \leq\ & \hat{x} \leq \lambda \end{aligned}$$

Poissonovo rozdělení se obvykle užívá pro vyjádření pravděpodobnosti počtu nastoupení sledovaného jevu v určitém časovém intervalu (počet poruch, nehod, katastrof, zmetků apod.) s malou pravděpodobností výskytu.



Obrázek 2: Grafy pravděpodobnostní funkce hypergeometrického rozdělení pro $N = 100$, $n = 10$ a různé hodnoty M. Přerušovaná čára je použita pouze pro odlišení jednotlivých pravděpodobnostních funkcí.

13. Poznámka Na Obrázku 3 jsou uvedeny příklady pravděpodobnostních funkcí Poissonova rozdělení.



Obrázek 3: Grafy pravděpodobnostní funkce Poissonova rozdělení pro různé hodnoty λ. Přerušovaná čára je použita pouze pro odlišení jednotlivých pravděpodobnostních funkcí.

14. Příklad Statistickým průzkumem bylo zjištěno, že během jedné minuty navštíví prodejnu průměrně 3 zákazníci. Najděte vhodný typ rozdělení pravděpodobnosti náhodné veličiny X vyjadřující počet zákazníků, kteří navštíví prodejnu během jedné minuty. Určete její pravděpodobnostní funkci $p(x)$, střední počet zákazníků $\mathrm{E}(X)$, rozptyl $\mathrm{D}(X)$ a směrodatnou odchylku $\sigma(X)$ počtu zákazníků, koeficient šikmosti $\mathrm{A}_3(X)$ a nejpravděpodobnější počet zákazníků za jednu minutu. Určete dále pravděpodobnost, že během jedné minuty přijde a) právě 1 zákazník, b) aspoň 1 zákazník, c) medián $x_{0.5}$ počtu zákazníků.

Řešení Nahradíme střední počet zákazníků, kteří navštíví prodejnu během jedné minuty, jejich průměrným počtem, tj. položíme $\mathrm{E}(X) = \bar{x}$. Vzhledem k tomu, že nemáme další informace o rozdělení pravděpodobnosti náhodné veličiny X (např. o rozptylu $\mathrm{D}(X)$ a koeficientu šikmosti $\mathrm{A}_3(X)$), použijeme Poissonovo rozdělení pravděpodobnosti $\mathrm{Po}(\lambda)$ s pravděpodobnostní funkcí

$$p(x) = \frac{3^x}{x!}e^{-3}, \quad x = 0, 1, \ldots$$

Ze vzorců můžeme vypočíst následující číselné charakteristiky.

Střední hodnota	$\mathrm{E}(X) = \lambda = 3,$
Rozptyl	$\mathrm{D}(X) = \lambda = 3,$
Směrodatná odchylka	$\sigma(X) = \sqrt{\mathrm{D}(X)} = \sqrt{3} \doteq 1.73205,$
Koeficient šikmosti	$\mathrm{A}_3(X) = 1/\sqrt{\lambda} = 1/\sqrt{3} \doteq 0.57735,$
Modus (nejpravděpodobnější počet zákazníků)	$\hat{x} = 2$ a 3, neboť $\lambda - 1 \le \hat{x} \le \lambda,$

Z pravděpodobnostní funkce pak lze přímo vypočíst pravděpodobnosti

a) $P(X = 1) = p(1) = \frac{3^1}{1!}e^{-3} \doteq 0.14936,$

b) $P(X \ge 1) = p(1) + p(2) + \ldots = 1 - p(0) = 1 - \frac{3^0}{0!}e^{-3} \doteq 1 - 0.04979 = 0.95021,$

c) Medián je $x_{0.5} = 3$, neboť $p(0)+p(1)+p(2) \doteq 0.42319 < 0.5$ a $p(0)+p(1)+p(2)+p(3) \doteq 0.64723 > 0.5$.

Spojitá rozdělení pravděpodobnosti

15. Pojem Náhodná veličina s **Rovnoměrným rozdělením** $\mathrm{R}(a, b)$, kde a, b jsou reálná čísla, $a < b$, má hustotu

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in \langle a, b\rangle \\ 0 & \text{pro } x \notin \langle a, b\rangle \end{cases}$$

16. Vlastnosti Náhodná veličina s Rovnoměrným rozdělením má distribuční funkci

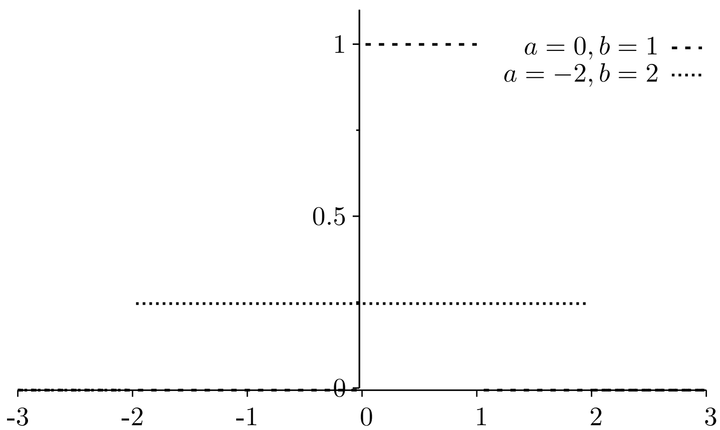
$$F(x) = \begin{cases} 0 & \text{pro } x \notin (-\infty, a) \\ \frac{x-a}{b-a} & \text{pro } x \in \langle a, b \rangle \\ 1 & \text{pro } x \notin (b, \infty) \end{cases}$$

a následující číselné charakteristiky

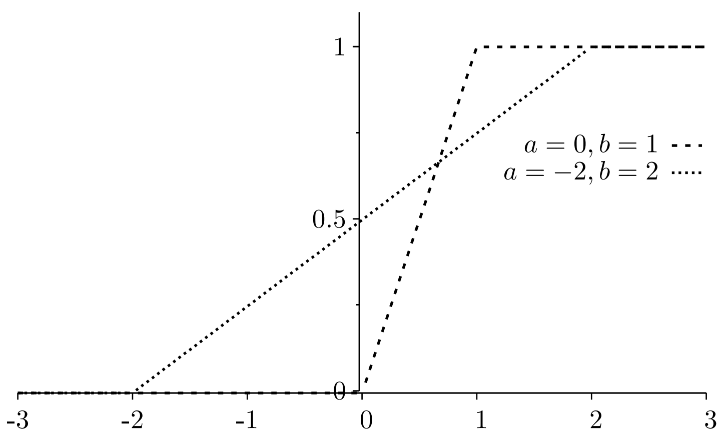
$$\begin{aligned} \mathrm{E}(X) &= x_{0.5} = \frac{a+b}{2}, \\ \mathrm{D}(X) &= \frac{(b-2)^2}{12}, \\ \mathrm{A}_3(X) &= 0 \end{aligned}$$

Rovnoměrné rozdělení slouží především k simulaci reálných procesů nebo numerickým výpočtům tzv. **metodou Monte Carlo** na počítači pomocí generátorů tzv. **pseudonáhodných čísel**.

17. Poznámka Na Obrázku 4 jsou grafy hustot pravděpodobnosti a na Obrázku 5 grafy odpovídajících distribučních funkcí rovnoměrného rozdělení pro různé hodnoty parametrů aa b.



Obrázek 4: Grafy hustot rovnoměrného rozdělení pro různé hodnoty parametrů a a b.



Obrázek 5: Grafy distribučních funkcí rovnoměrného rozdělení pro různé hodnoty parametrů a a b.

18. Příklad K přerušení optického kabelu v délce 500 m může dojít v libovolné vzdálenosti od jeho počátku, přičemž pravděpodobnost náhodného jevu, že dojde k přerušení v nějakém úseku je přímo úměrná délce úseku a nezávisí na jeho poloze. Určete rozdělení pravděpodobnosti náhodné veličiny X vyjadřující vzdálenost místa přerušení od počátku, její hustotu pravděpodobnosti a základní číselné charakteristiky a pravděpodobnost, že k přerušení kabelu dojde v úseku od 300 m do 400 m.

Řešení Náhodná veličina X má rozdělení $\mathrm{R}(a,b)$, kde $a = 0$ a b=500 s hustotou pravděpodobnosti

$$f(x) = \begin{cases} \frac{1}{500} & \text{pro } x \in \langle 0, 500 \rangle \\ 0 & \text{pro } x \notin \langle 0, 500 \rangle \end{cases}$$

Ze vzorců můžeme vypočíst následující číselné charakteristiky.

Střední hodnota a medián	$\mathrm{E}(X) = x_{0.5} = \frac{0+500}{2} = 250\mathrm{m}$,
Rozptyl	$\mathrm{D}(X) = \frac{(500-0)^2}{12} \doteq 20833.33\mathrm{m}^2$,
Směrodatná odchylka	$\sigma(X) = \sqrt{\mathrm{D}(X)} \doteq \sqrt{20833.3} \doteq 144.34\mathrm{m}$,
Koeficient šikmosti	$\mathrm{A}_3(X) = 0$.

Z pravděpodobnostní funkce pak lze přímo vypočíst pravděpodobnost

$$\mathrm{P}(300 \leq X \leq 400) = F(400) - F(300) = \frac{400}{500} - \frac{300}{500} = 0.2.$$

19. Pojem Náhodná veličina s **Normálním rozdělením** $\mathrm{N}(\mu, \sigma^2)$, kde μ, σ^2 jsou reálná čísla, a $\sigma^2 > 0$, má hustotu

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}, \quad x \in (-\infty, \infty)$$

20. Vlastnosti Náhodná veličina s Normálním rozdělením má následující číselné charakteristiky

$$\begin{aligned} \mathrm{E}(X) &= x_{0.5} = \hat{x} = \mu, \\ \mathrm{D}(X) &= \sigma^2, \\ \mathrm{A}_3(X) &= 0 \end{aligned}$$

Normální rozdělení, které je nejčastěji užívaným rozdělením, je také nazýváno **Gaussovo rozdělení**. Má řadu významných teoretických vlastností a z hlediska aplikací bývá vhodné k vyjádření náhodných veličin, které lze interpretovat jako aditivní výsledek mnoha nezávislých vlivů (např. chyba měření, odchylka rozměru výrobku od požadované hodnoty apod.).

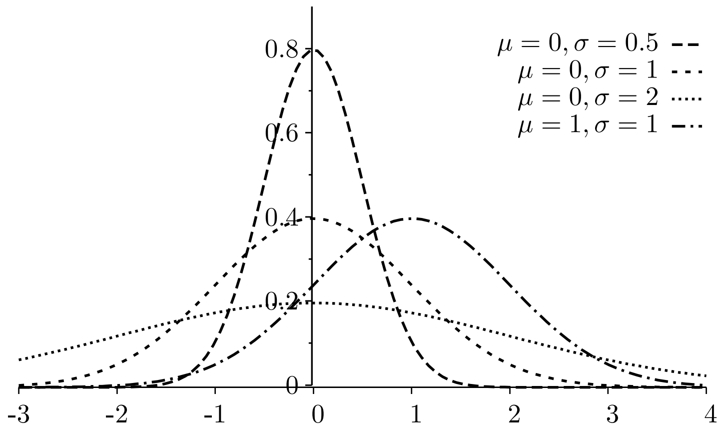
21. Poznámka Na Obrázku 6 jsou grafy hustot pravděpodobnosti a na Obrázku 7 grafy odpovídajících distribučních funkcí normálního rozdělení pro různé hodnoty parametrů μ a σ.

22. Vlastnosti Jestliže náhodná veličina X má normální rozdělení $\mathrm{N}(\mu, \sigma^2)$, pak náhodná veličina $Y = aX + b$, kde a, b jsou reálná čísla, $a \neq 0$, má normální rozdělení $\mathrm{N}(a\mu + b, a^2\sigma^2)$.

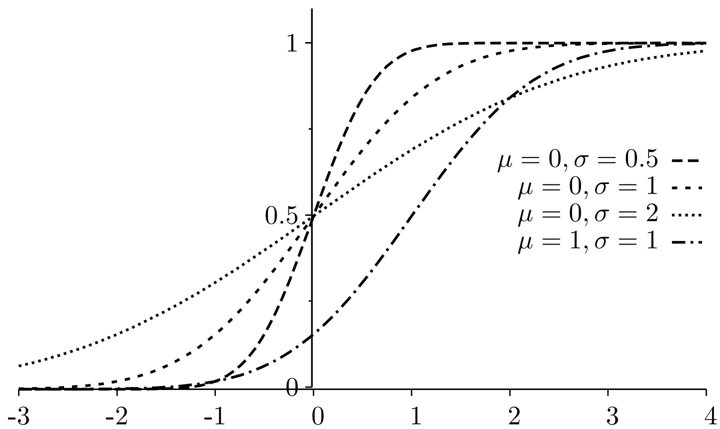
23. Pojem Transformací náhodné veličiny X s normálním rozdělením $\mathrm{N}(\mu, \sigma^2)$ na náhodnou veličinu

$$U = \frac{X - \mu}{\sigma}$$

dostaneme náhodnou veličinu s **normovaným (základním) normálním rozdělením** $\mathrm{N}(0,1)$ s distribuční funkcí $\Phi(u)$.



Obrázek 6: Grafy hustot Normálního rozdělení s různými hodnotami parametrů μ, σ^2



Obrázek 7: Grafy distribučních funkcí Normálního rozdělení s různými hodnotami parametrů μ, σ^2

24. Vlastnosti Pro hodnoty $\Phi(u)$ distribuční funkce normovaného normálního rozdělení platí

$$\Phi(-u) = 1 - \Phi(u).$$

Pro kvantily normovaného normálního rozdělení je

$$u_{1-P} = -u_P, \quad 0 < P < 1.$$

Jestliže náhodná veličin X má normální rozdělení $\mathrm{N}(\mu, \sigma^2)$, potom lze její distribuční funkci vyjádřit jako

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

a její kvantily jsou

$$x_P = \mu + \sigma u_P, \quad 0 < P < 1.$$

Hodnoty distribuční funkce $\Phi(u)$ normované náhodné veličiny U jsou tabelovány v tabulce **T1**. K výpočtu hodnot $\Phi(u)$ a kvantilů u_P na PC lze také použít vhodný software (např. Statistica, Statgraphics, Excel aj.).

25. Příklad Určete pravděpodobnost, že náhodná veličina X s normálním rozdělením pravděpodobnosti N(20;16), nabude hodnotu

1. menší než 16,
2. větší než 20,
3. v mezích od 12 do 28,
4. menší než 12 nebo větší než 28.

Řešení Ze vztahu $F(x) = \Phi\left(\frac{x-20}{4}\right)$ a tabulky **T1** dostaneme:

1. $P(X < 16) = F(16) = \Phi((16-20)/4) = \Phi(-1) = 1 - \Phi(1) = 1 - 0.84135 = 0.15865$,
2. $P(X > 20) = 1 - P(X \leq 20) = 1 - F(20) = 1 - \Phi((20-20)/4) = 1 - \Phi(0) = 1 - 0.5 = 0.5$,
3. $P(12 \leq X \leq 28) = F(28) - F(12) = \Phi((28-20)/4) - \Phi((12-20)/4) = \Phi(2) - \Phi(-2) =$
$= \Phi(2) - (1 - \Phi(2)) = 2\Phi(2) - 1 = 2 \cdot 0.97725 - 1 = 0.95450$,
4. $P((X < 12) \vee (X > 28)) = 1 - P(12 \leq X \leq 28) = 1 - 0.9545 = 0.0455$.

26. Příklad Měření délkového rozměru je zatíženo systematickou chybou 0.5 mm a náhodnou chybou s normálním rozdělením pravděpodobnosti s rozptylem 0.09 mm^2. Určete, pro jakou hodnotu δ bude celková chyba jednoho měření v mezích $0.5 - \delta$ až $0,5 + \delta$ s pravděpodobností 0.95.

Řešení Chyba jednoho měření X má normální rozdělení s parametry $\mu = 0,5$ a $\sigma^2 = 0.09$, neboť u náhodné chyby předpokládáme, že má nulovou střední hodnotu, takže

$$\begin{aligned} P(0.5 - \delta \leq X \leq 0.5 + \delta) &= F(0.5 + \delta) - F(0.5 - \delta) = \\ &= \Phi\left(\frac{\delta}{0.3}\right) - \Phi\left(-\frac{\delta}{0.3}\right) = 2\Phi\left(\frac{\delta}{0.3}\right) - 1 = 0.95. \end{aligned}$$

Odtud je $\Phi(\frac{\delta}{0.3}) = 0.975$, takže $\frac{\delta}{0.3} = u_{0.975}$. Protože z tabulky **T1** je $u_{0.975} = 1.960$, je $\delta = 0.3 \cdot 1.960 = 0.588$. S pravděpodobností 0.95 bude celková chyba jednoho měření v intervalu $\langle -0.088; 1.088 \rangle$mm.