MUNI
ECON

Spojité náhodné veličiny a jejich pravděpodobnostní rozdělení

Povinná literatura: Mann (2016), Kapitola 6

Z čeho studovat šestou lekci?

Povinná literatura:	Mann (2016), kap. 06
Příprava na cvičení:	Leaflet 07 Koncepty a procedury, cv. 07, kap. 06
Příprava na zkoušku:	Mann (2016), kap. 06 Leaflet 07 Sbírka úloh, kap. 06 Koncepty a procedury, cv. 07, kap. 06

Motivační vstup - využití normálního rozdělení

- **Modelování výnosů z investic:** Odhad rozložení výnosů akcií, dluhopisů a jiných finančních nástrojů.
- **Posuzování rizik a nejistot:** Predikce rizik spojených s investicemi a výkyvy na finančních trzích.
- **Analýza spotřebitelských preferencí:** Studium spotřebitelského chování a rozložení výdajů mezi různými skupinami populace.
- **Statistické testování ekonomických hypotéz:** Ověřování ekonomických teorií pomocí analýzy dat a normálních rozdělení.
- **Makroekonomická data:** Analýza distribuce makroekonomických ukazatelů, jako je inflace, nezaměstnanost či HDP.
- **Finanční plánování a analýza:** Predikce rozpočtů, tržeb a výdajů na základě historických dat a jejich rozložení.

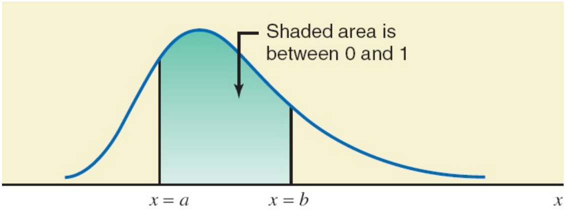
Spojitá pravděpodobnostní rozdělení

Dvě vlastnosti:

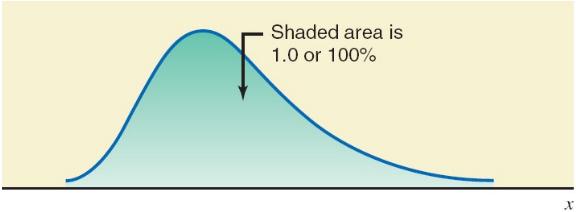
1. Pravděpodobnost, že náhodná veličina x nabývá hodnoty v jakémkoli intervalu, je v rozmezí 0 až 1.
2. Celková pravděpodobnost všech (neslučitelných) intervalů, ve kterých může náhodná veličina x nabývat hodnoty, je 1.0.

Plocha pod křivkou rozdělení

Obrázek: Plocha mezi dvěma body

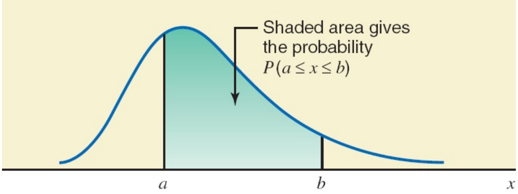


Obrázek: Plocha pod celou křivkou

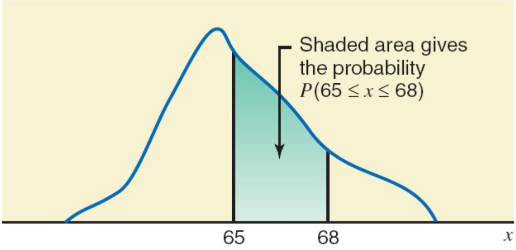


Plocha pod křivkou rozdělení

Obrázek: Plocha mezi dvěma body vyjádřená jako pravděpodobnost

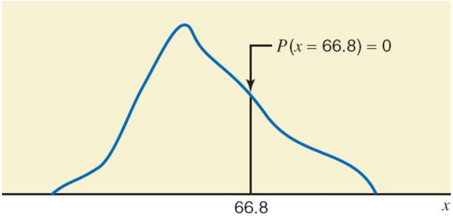


Obrázek: Pravděpodobnost, že x leží od 65 do 68

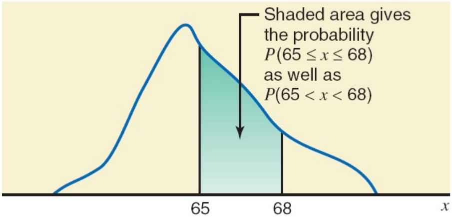


Plocha pod křivkou rozdělení

Obrázek: Pravděpodobnost jedné konkrétní hodnoty x je nulová



Obrázek: Pravděpodobnost, že x leží od 65 do 68 a mezi 65 a 68 je stejná



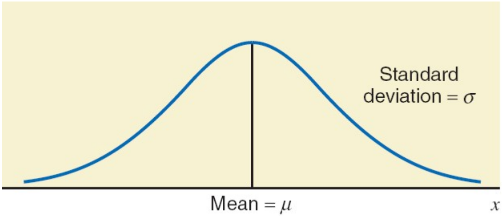
Normální rozdělení

Normální rozdělení pravděpodobnosti $N(\mu, \sigma^2)$ v grafickém zobrazení vytváří zvonovitou křivku (Gaussovu křivku), pro kterou platí:

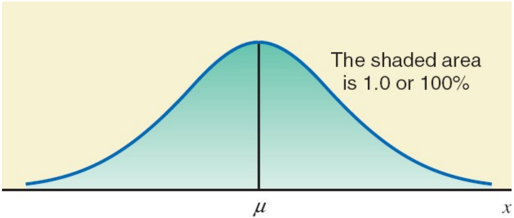
1. Celková plocha pod křivkou je 1.0.
2. Křivka je symetrická vzhledem k střední hodnotě (průměru).
3. Oba chvosty (konce) křivky se nekonečně rozšiřují.

Normální rozdělení

Obrázek: Normální rozdělení pravděpodobnosti $N(\mu, \sigma^2)$

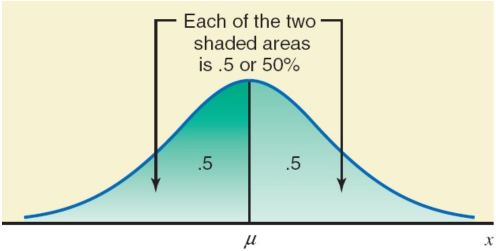


Obrázek: Celková plocha pod křivkou

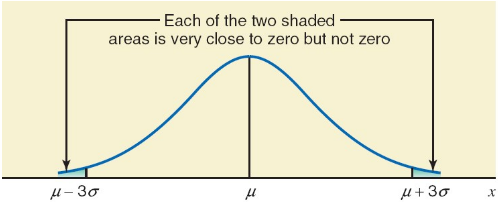


Normální rozdělení

Obrázek: Normální rozdělení $N(\mu, \sigma^2)$ je symetrické

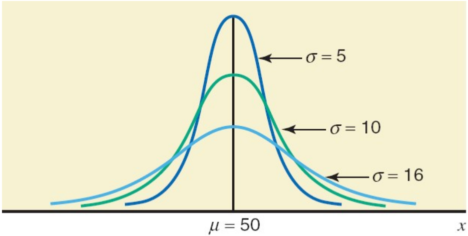


Obrázek: Konce rozdělení $N(\mu, \sigma^2)$

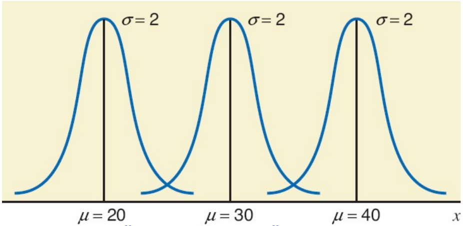


Normální rozdělení

Obrázek: Tři normální rozdělení se stejným μ a rozdílným σ^2

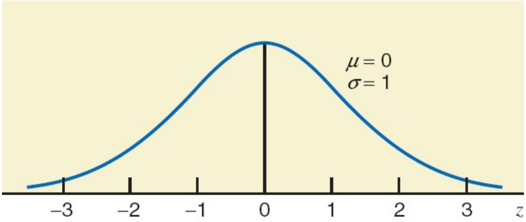


Obrázek: Tři normální rozdělení s rozdílným μ a stejným σ^2



Standardizované normální rozdělení

Normální rozdělení se střední hodnotou $\mu = 0$ a směrodatnou odchylkou $\sigma = 1$ se nazývá **standardizované normální rozdělení**.



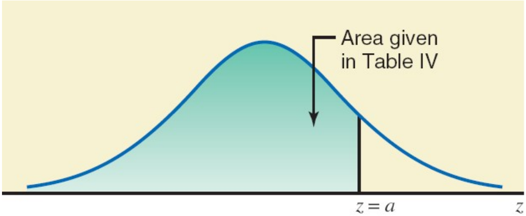
Z-hodnota nebo Z-skóre

Jednotky označené na horizontální ose standardizované křivky normálního rozdělení jsou vyjádřeny pomocí **z** a nazývají se ***z*-hodnoty nebo *z*-skóre**.

Konkrétní hodnota **z** udává vzdálenost mezi průměrem a bodem reprezentovaným hodnotou **z** ve směrodatných odchylkách.

Příklad 1

Zadání: Určete obsah plochy pod křivkou standardizovaného normálního rozdělení nalevo od $z = a = 1.95$.



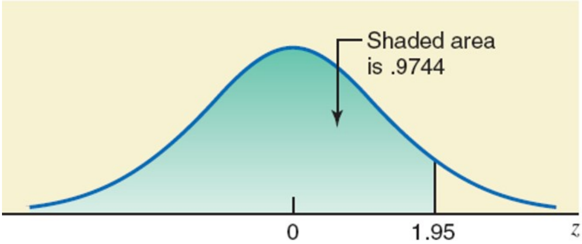
Příklad 1: Řešení

Table 6.2 Area Under the Standard Normal Curve to the Left of $z = 1.95$

z	.00	.01	...	$\boxed{.05}$	...	.09
−3.4	.0003	.0003	...	.0003	...	.0002
−3.3	.0005	.0005	...	.0004	...	.0003
−3.2	.0007	.0007	...	.0006	...	.0005
.	.	.	...	.	...	.
.	.	.	...	.	...	.
.	.	.	...	.	...	.
$\boxed{1.9}$	.9713	.9719	...	$\boxed{.9744}$	...	.9767
.	.	.	...	.	...	.
.	.	.	...	.	...	.
.	.	.	...	.	...	.
3.4	.9997	.9997	...	.9997	...	.9998

Required area

Příklad 1: Řešení



Příklad 2

Zadání: Určete obsah plochy pod křivkou standardizovaného normálního rozdělení od z = -2.17 do z = 0.

Řešení:
Abychom našli obsah plochy od $z = -2.17$ do $z = 0$, nejprve určíme obsahy vlevo od $z = 0$ a vlevo od $z = -2.17$ např. v tabulce IV.

Jak je uvedeno v tabulce na následujícím slidu, tyto dva obsahy jsou .5 a .0150. Hledaný obsah získáme odečtením .0150 od .5.

Obsah plochy od -2.17 do 0 lze zapsat jako
$P(-2.17 \leq z \leq 0) = .5000 - .0150 = \mathbf{.4850}$

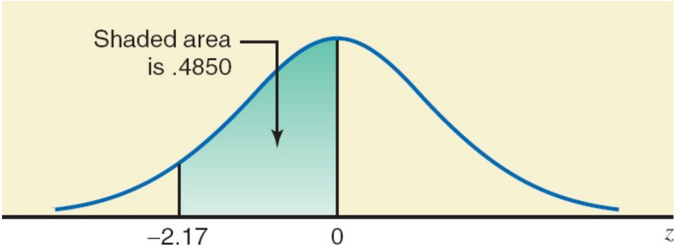
Příklad 2: Řešení

z	$\boxed{00}$	.01	...	$\boxed{.07}$	...	.09
−3.4	.0003	.0003	...	.0003	...	.0002
−3.3	.0005	.0005	...	.0004	...	.0003
−3.2	.0007	.0007	...	.0005	...	.0005
.	.	.	...	.	...	.
.	.	.	...	.	...	.
$\boxed{-2.1}$	.0179	.0174	...	$\boxed{.0150}$	...	.0143
.	.	.	...	.	...	.
.	.	.	...	.	...	.
$\boxed{0.0}$	$\boxed{.5000}$	.5040	...	.5279	...	.5359
.	.	.	...	.	...	.
.	.	.	...	.	...	.
3.4	.9997	.9997	...	.9997	...	.9998

Area to the left of $z = 0$

Area to the left of $z = -2.17$

Příklad 2: Řešení graficky



Příklad 3

Zadání: Určete následující pravděpodobnosti (plochy pod křivkou) pro standardizované normální rozdělení.

a) $P(-1.56 < z < 2.31)$
b) $P(z > -0.75)$
c) $P(z < -5.35)$

Řešení:

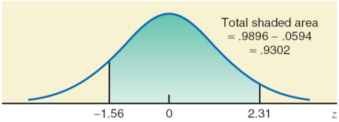
a)
$$\begin{aligned} P(-1.56 < z < 2.31) &= \text{Oblast mezi -1.56 a 2.31} \\ &= .9896 - .0594 = \mathbf{.9302} \end{aligned}$$

b)
$$\begin{aligned} P(z > -0.75) &= \text{Oblast vpravo od} - 0.75 \\ &= 1.00 - .2266 = \mathbf{.7734} \end{aligned}$$

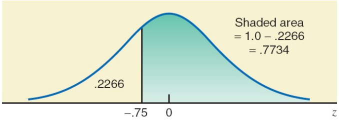
c)
$$\begin{aligned} P(z < -5.35) &= \text{Oblast vlevo od} - 5.35 \\ &= \mathbf{0.00} \text{ přibližně} \end{aligned}$$

Příklad 3: Řešení graficky

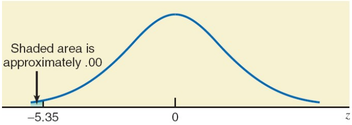
Obrázek: Řešení a)



Obrázek: Řešení b)

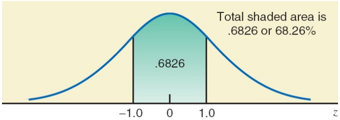


Obrázek: Řešení c)

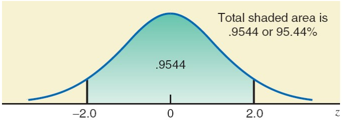


Plocha v okolí střední hodnoty

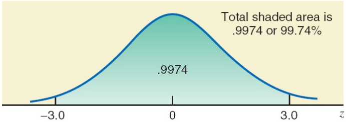
Obrázek: Okolí: $\pm$ 1x směrodatná odchylka



Obrázek: Okolí: $\pm$ 2x směrodatná odchylka



Obrázek: Okolí: $\pm$ 3x směrodatná odchylka



Spojité NV a jejich pravděpodobnostní rozdělení

- Spojitá pravděpodobnostní rozdělení
- Normální rozdělení
- Standardizované normální rozdělení

Standardizované a normální rozdělení

Aplikace normálního rozdělení

Stanovení hodnot z a x při známé ploše pod křivkou NR

Aproximace binomického rozdělení normálním rozdělením

Standardizované a normální rozdělení

Převod náhodné veličiny *x* na náhodnou veličinu *z*

Pro náhodnou veličinu *x* z normálního rozdělení, lze určitou hodnotu *x* převést na odpovídající hodnotu **z** použitím vzorce:

$$\mathbf{z} = \frac{x - \mu}{\sigma}$$

kde μ a σ jsou střední hodnota (průměr) a směrodatná odchylka normálního rozdělení náhodné veličiny *x* .

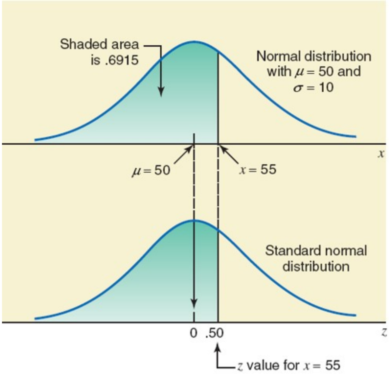
Příklad 4

Zadání: Nechť x je spojitá náhodná veličina, která má normální rozdělení se střední hodnotou 50 a směrodatnou odchylkou 10. Převeďte následující hodnoty x na hodnoty z a najděte pravděpodobnost nalevo od bodu $x = 55$.

Řešení:

$$z = \frac{x - \mu}{\sigma} = \frac{55 - 50}{10} = .50$$

$$P(x < 55) = P(z < .50) = \mathbf{.6915}$$



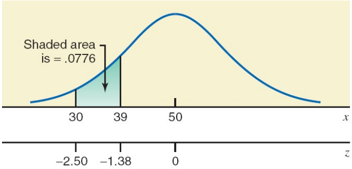
Příklad 5

Zadání: Nechť x je spojitá náhodná veličina řídící se normálním rozdělením s parametry $\mu = 50$ a $\sigma = 8$. Určete $P(30 \leq x \leq 39)$.

Řešení:

Pro $x = 30$: $z = \frac{30-50}{8} = -2.50$

Pro $x = 39$: $z = \frac{39-50}{8} = -1.38$



$$P(30 \leq x \leq 39) = P(-2.50 \leq z \leq -1.38) = .0838 - .0062 = \mathbf{.0776}$$

Aplikace normálního rozdělení

V předcházejících oddílech jsme se zabývali normálním rozdělením, způsoby převodu normálního rozdělení na standardizované normální rozdělení a metodami pro určení obsahu pod křivkou normálního rozdělení.

V této části jsou uvedeny příklady, které demonstrují použití normálního rozdělení.

Příklad 6

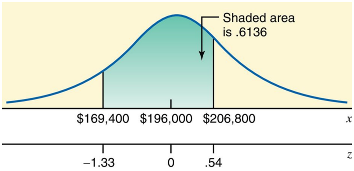
Zadání: Dle *Zprávy o kompenzaci lékařů za rok 2015* si američtí internisté v roce 2014 vydělali v průměru 196 000 dolarů.

Předpokládejme, že příjmy všech amerických internistů za rok 2014 mají normální rozdělení se střední hodnotou 196 000 dolarů a směrodatnou odchylkou 20 000 dolarů.

Určete pravděpodobnost, že příjmy náhodně vybraného amerického internisty za rok 2014 jsou mezi 169 400 dolarů a 206 800 dolarů.

Příklad 6: Řešení

Graficky:



Pro $x = 169\,400$:

$$z = \frac{169\,400 - 196\,000}{20\,000} = -1.33$$

Pro $x = 206\,800$:

$$z = \frac{206\,800 - 196\,000}{20\,000} = .54$$

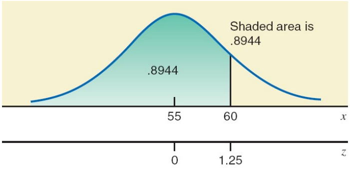
$$\begin{aligned} P(169\,400 \leq x \leq 206\,800) &= P(-1.33 \leq z \leq 0.54) \\ &= 0.7054 - 0.0918 \\ &= \mathbf{0.6136} = \mathbf{61.36}\% \end{aligned}$$

Příklad 7

Zadání: Závodní automobil je jednou z mnoha hraček, které vyrábí společnost Mack Corporation. Doba montáže této hračky se řídí normálním rozdělením se střední hodnotou 55 minut a směrodatnou odchylkou 4 minuty. Firma ukončuje svou činnost vždy v 17:00. Pokud pracovník začne sestavovat závodní automobil v 16:00, jaká je pravděpodobnost, že práci stihne dokončit před koncem pracovní doby?

Příklad 7: Řešení

Graficky:



Pro $x = 60$: $z = \frac{60-55}{4} = 1.25$

$P(x \leq 60) = P(z \leq 1.25) = \mathbf{.8944}$

Pravděpodobnost, že tento pracovník dokončí sestavení závodního auta před uzavřením firmy, je **.8944**.

Příklad 8

Zadání: Společnost Bublinky s.r.o vyrábí mnoho druhů nápojů, včetně Oranžové limonády. Plnicí stroje jsou nastaveny tak, aby do každé 12 uncové plechovky Oranžové limonády nalily 12 uncí sodovky.

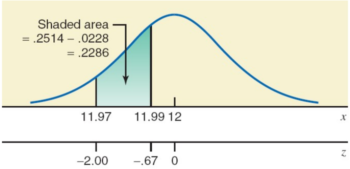
Avšak skutečné množství nalité sodovky do každé plechovky není přesně 12 uncí; liší se od plechovky k plechovce. Bylo zjištěno, že čisté množství sodovky v takové plechovce má normální rozdělení se střední hodnotou 12 uncí a směrodatnou odchylkou 0.015 unce.

(a) Jaká je pravděpodobnost, že náhodně vybraná plechovka Oranžové limonády obsahuje mezi 11.97 a 11.99 uncí sodovky?

(b) Kolik procent plechovek Oranžové limonády obsahuje mezi 12.02 a 12.07 uncí sodovky?

Příklad 8: Řešení a)

Graficky:

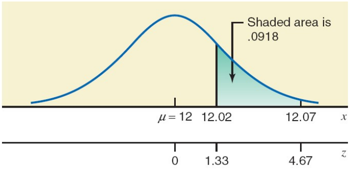


$$x = 11.97: \quad z = \frac{11.97-12}{0.015} = -2.00$$
$$x = 11.99: \quad z = \frac{11.99-12}{0.015} = -0.67$$

$$\begin{aligned} P(11.97 \leq x \leq 11.99) &= P(-2.00 \leq z \leq -0.67) \\ &= .2514 - .0228 \\ &= \mathbf{.2286} \end{aligned}$$

Příklad 8: Řešení b)

Graficky:



$$x = 12.02: \quad z = \frac{12.02 - 12}{0.015} = 1.33$$
$$x = 12.07: \quad z = \frac{12.07 - 12}{0.015} = 4.67$$

$$\begin{aligned} P(12.02 \leq x \leq 12.07) &= P(1.33 \leq z \leq 4.67) \\ &= 1.0 - .9082 \\ &= \mathbf{.0918} \end{aligned}$$

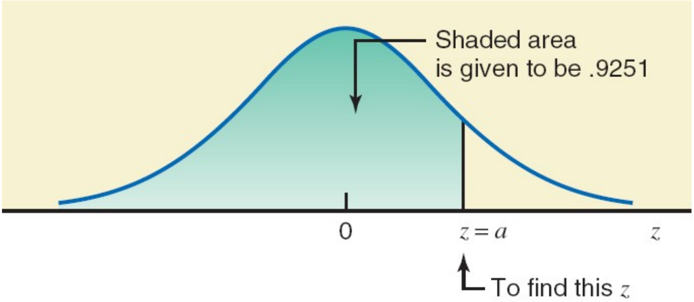
Stanovení hodnot z a x při známé ploše pod křivkou normálního rozdělení

Doposud jsme se zabývali způsobem, jak určit plochu pod křivkou normálního rozdělení pro určitý interval hodnot z nebo x.

Teď si ukážeme opačný postup a vysvětlíme si, jak zjistit příslušné hodnoty z nebo x, pokud je známa plocha pod křivkou normálního rozdělení.

Příklad 9

Zadání : Určete hodnotu z tak, aby obsah pod křivkou standardizovaného normálního rozdělení vlevo od z byl 0.9251.



Příklad 9: Řešení

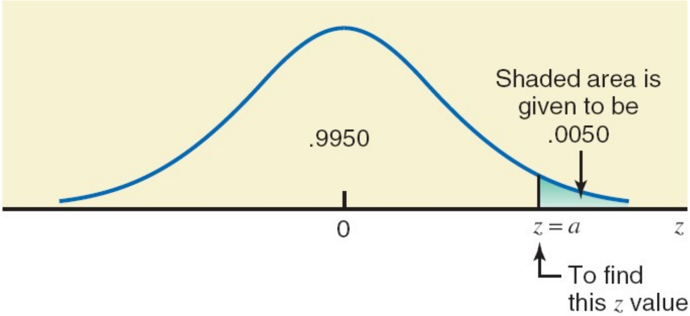
Table 6.4 Finding the z Value When Area Is Known

z	.00	.01	...	$\boxed{.04}$	...	.09
−3.4	.0003	.0003	...		...	.0002
−3.3	.0005	.0005	...		...	.0003
−3.2	.0007	.0007	...		...	.0005
·	·	·	...		...	·
·	·	·	...		...	·
·	·	·	...		...	·
$\boxed{1.4}$				$\boxed{.9251}$	...	...
·	·	·	...	·	...	·
·	·	·	...	·	...	·
·	·	·	...	·	...	·
3.4	.9997	.9997	...	.9997	...	.9998

We locate this value in Table IV of Appendix B

Příklad 10

Zadání : Určete hodnotu z tak, aby obsah pod křivkou standardizovaného normálního rozdělení v pravém chvostu (konci) byl 0.0050.



Příklad 10: Řešení

Pro nalezení hodnoty z, nejprve najdeme oblast nalevo od z tak, že

$$\text{Oblast nalevo od } z = 1.0 - .0050 = .9950$$

Nyní hledáme číslo .9950 v těle tabulky normálního rozdělení. Tabulka IV neobsahuje .9950. Takže najdeme hodnotu nejbližší k .9950, která je buď .9949 nebo .9951. Pokud zvolíme .9951, pak

$$z = 2.58$$

Pokud zvolíme .9949, pak

$$z = 2.57$$

Můžeme použít libovolnou z těchto dvou hodnot.

Hledání hodnoty x u normálního rozdělení

Pokud známe střední hodnotu μ a směrodatnou odchylku σ normálního rozdělení a máme zadanou oblast pod křivkou na levé straně bodu x, hodnotu x vypočteme pomocí vzorce

$$x = \mu + z\sigma$$

Příklad 11

Zadání : Je známo, že životnost kalkulačky vyrobené společností Calculators Corporation se řídí normálním rozdělením se střední hodnotou 54 měsíců a směrodatnou odchylkou 8 měsíců.

Jak dlouhá by měla být záruční doba na výměnu nefunkční kalkulačky, pokud společnost nechce vyměnit více než 1 % všech prodaných kalkulaček?

Příklad 11: Řešení

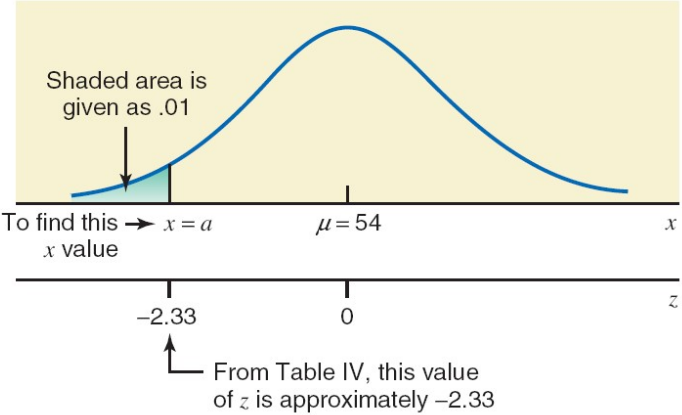
Nechť x je životnost kalkulačky. Oblast nalevo od x je .01 nebo 1%.

Najdeme hodnotu z z tabulky normálního rozdělení pro .0100. Tabulka IV neobsahuje hodnotu, která je přesně .0100 a tak je hodnota nejbližší k .0100 v tabulce .0099. Tedy z = -2.33.

$$x = \mu + z\sigma = 54 + (-2.33)(8) = 54 - 18.64 = \mathbf{35.36}$$

Společnost, aby nemusela vyměňovat více než 1 % kalkulaček, by tedy měla vyměňovat všechny kalkulačky, které začnou vykazovat poruchy do 35.36 měsíce (což lze zaokrouhlit na 35 měsíců) od data nákupu.

Příklad 11: Řešení graficky



Aproximace binomického rozdělení normálním rozdělením

1. Binomické rozdělení se aplikuje na diskrétní náhodnou veličinu.
2. Každé opakování, nazývané pokus, binomického experimentu vede k jednomu ze dvou možných výsledků (nebo jevů), a to buď úspěchu nebo neúspěchu.
3. Pravděpodobnosti obou možných výsledků (nebo jevů) zůstávají stejné pro každé opakování experimentu.
4. Pokusy jsou nezávislé.

Aproximace binomického rozdělení normálním rozdělením

Binomický vzorec, který udává pravděpodobnost x úspěchů v n pokusech, je

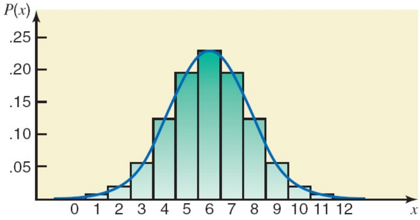
$$P(x) = \binom{n}{x} p^x q^{n-x}.$$

Normální rozdělení se používá jako aproximace binomického rozdělení, když jsou obě hodnoty np a nq větší než 5, tzn. když

$$np > 5 \quad \text{a} \quad nq > 5.$$

Aproximace binomického rozdělení normálním rozdělením

Ukázka : n = 12 a p = 0.50
Binomické rozdělení pravděpodobnosti



x	$P(x)$
0	.0002
1	.0029
2	.0161
3	.0537
4	.1208
5	.1934
6	.2256
7	.1934
8	.1208
9	.0537
10	.0161
11	.0029
12	.0002

Příklad 12

Zadání: Podle odhadu má 50 % lidí v USA alespoň jednu kreditní kartu. Jaká je pravděpodobnost, že z náhodného vzorku 30 osob bude mít alespoň jednu kreditní kartu 19 lidí?

Řešení (přesné):

$$n = 30, \quad p = 0.50, \quad q = 1 - p = 0.50$$

$$x = 19, \quad n - x = 30 - 19 = 11$$

Podle binomického vzorce je přesná pravděpodobnost

$$P(19) = \binom{30}{19}(0.5)^{19}(0.5)^{11} = 0.0509$$

Příklad 12: Řešení aproximací

Zadání (rozšířeno): Vyřešme tento problém s využitím normálního rozdělení jako aproximací binomického rozdělení.

Prvně ověříme, zda lze vůbec aproximaci provést

$$np = 30(0.50) = 15 > 5,$$

$$nq = 30(0.50) = 15 > 5.$$

Jelikož obě hodnoty podmínek jsou větší než 5, můžeme použít normální rozdělení jako aproximaci k vyřešení tohoto binomického problému.

Příklad 12: Řešení aproximací

Krok 1. Vypočítejte μ a σ pro binomické rozdělení.

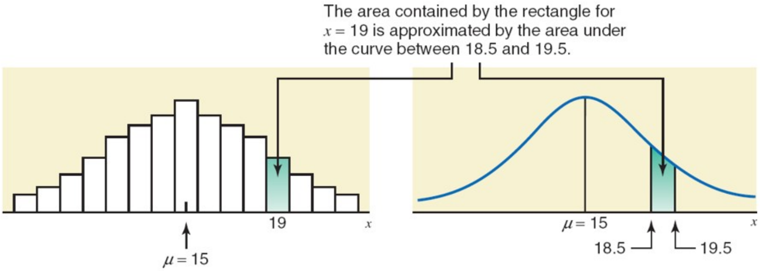
$$\mu = np = 30(0.50) = 15$$

$$\sigma = \sqrt{npq} = \sqrt{30(0.50)(0.50)} = 2.73861279$$

Krok 2. Převeďte diskrétní náhodnou veličinu na spojitou náhodnou náhodou veličinu (provedením *opravy na spojitost*).

Vsuvka - Oprava na spojitost

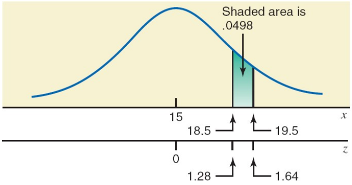
Faktorem opravy na spojitost chápeme přičtení 0.5 a/nebo odečtení 0.5 od hodnot(y) x při použití normálního rozdělení jako aproximace binomického rozdělení, kde x představuje počet úspěchů v n počtu pokusů.



Příklad 12: Řešení aproximací

Krok 3. Spočítejte požadovanou pravděpodobnost pomocí normálního rozdělení.

Graficky:



Pro $x = 18.5$:

$$z = \frac{18.5 - 15}{2.73861279} = 1.28$$

Pro $x = 19.5$:

$$z = \frac{19.5 - 15}{2.73861279} = 1.64$$

$$P(18.5 \leq x \leq 19.5) = P(1.28 \leq z \leq 1.64) = 0.9495 - 0.8997 = 0.0498$$

Příklad 12: Řešení aproximací

Na základě aproximace normálním rozdělením je pravděpodobnost, že bude mít alespoň jednu kreditní kartu 19 osob ze vzorku 30, **přibližně 0.0498**.

Použitím binomického vzorce získáme **přesnou pravděpodobnost 0.0509**.

Chyba způsobená použitím aproximace normálním rozdělením je 0.0509 - 0.0498 = 0.0011.

Tedy přesná pravděpodobnost je podhodnocena o 0.0011, pokud se použije aproximace normálním rozdělením.

Příklad 13

Zadání: Podle průzkumu 32 % lidí pracujících z domova uvedlo, že největší výhodou práce z domova je absence dojíždění. Předpokládejme, že tento výsledek platí pro celou současnou populaci lidí pracujících z domova.

Jaká je pravděpodobnost, že v náhodném vzorku 400 lidí pracujících z domova 108 až 122 lidí uvede, že největší výhodou práce z domova je, že nemusí dojíždět?

Příklad 13: Řešení aproximací

Krok 1. $n = 400, \quad p = .32, \quad q = 1 - .32 = .68$

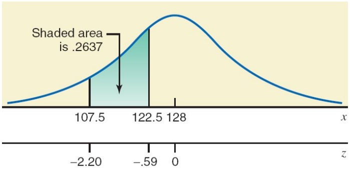
$$\mu = np = 400(.32) = 128$$
$$\sigma = \sqrt{npq} = \sqrt{400(.32)(.68)} = 9.32952303$$

Krok 2.

$x = 107.5 \colon z = \frac{107.5 - 128}{9.32952303} = -2.20$

$x = 122.5 \colon z = \frac{122.5 - 128}{9.32952303} = -0.59$

Graficky:



Krok 3.

$$P(107.5 \leq x \leq 122.5) = P(-2.20 \leq z \leq -0.59)$$
$$= .2776 - .0139 = .2637$$

Příklad 14

Zadání: Podle průzkumu mezi americkými dospělými 61 % žen podporuje umístění semaforů na křižovatkách. Předpokládejme, že tento procentní podíl platí pro současnou populaci amerických dospělých žen.

Jaká je pravděpodobnost, že 500 nebo více takových žen v náhodném vzorku 800 žen bude podporovat umístění semaforů na křižovatkách?

Příklad 14: Řešení aproximací

Krok 1. $n = 800,\ p = .61,\ q = 1 - .61 = .39$

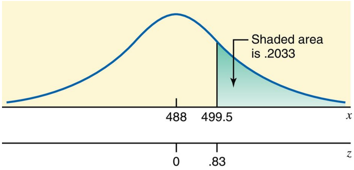
$$\mu = np = 800(.61) = 488$$

$$\sigma = \sqrt{npq} = \sqrt{800(.61)(.39)} = 13.79565149$$

Krok 2.

$x = 499.5$: $z = \frac{499.5-488}{13.79565149} = .83$

Graficky:



Krok 3.

$$P(x \geq 499.5) = P(z \geq .83) = 1.0 - .7967 = .2033$$

Pravděpodobnost, že 500 nebo více dospělých amerických žen z náhodného vzorku 800 žen bude podporovat umístění semaforů na křižovatkách, je přibližně .2033.

Shrnutí přednášky:

Spojité NV a jejich pravděpodobnostní rozdělení
- Spojitá pravděpodobnostní rozdělení
- Normální rozdělení
- Standardizované normální rozdělení

Standardizované a normální rozdělení

Aplikace normálního rozdělení

Stanovení hodnot z a x při známé ploše pod křivkou NR

Aproximace binomického rozdělení normálním rozdělením

Co si nastudovat na přespříští týden?

Příprava na cvičení: Leaflet 07
Koncepty a procedury, cv. 07, kap. 06

Povinná literatura: Mann (2016), Kapitola 7

Děkuji za pozornost!