5. cvičení: Úlohy o jednom náhodném výběru z normálního a alternativního rozložení.

Příklad 1.: Lze předpokládat, že hmotnost pomerančů dodávaných do obchodní sítě se řídí normálním rozložením se střední hodnotou 170 g a směrodatnou odchylkou 12 g. Jaká je pravděpodobnost, že celková hmotnost 9 náhodně vybraných pomerančů balených do síťky překročí 1,5 kg?
Výsledek: 0,797.

Výpočet pomocí systému STATISTICA
Uvědomíme si, že výběrový úhrn $\sum_{i=1}^{9} X_i$ se řídí normálním rozložením se střední hodnotou $9 \cdot 170 = 1530$ a rozptylem $9 \cdot 144 = 1296$.
Vytvoříme datový soubor o jedné proměnné a o jednom případu. Do Dlouhého jména této proměnné napíšeme
=1 – INormal(1500;1530;sqrt(1296)). Zjistíme, že hledaná pravděpodobnost je 0,797.
(Funkce INormal(x;μ;σ) počítá hodnotu distribuční funkce rozložení $N(\mu,\sigma^2)$ v bodě x.)

Příklad 2.: Počet bodů v testu inteligence je náhodná veličina, která se řídí rozložením N(100,225). Jaká je pravděpodobnost, že průměr v náhodně vybrané skupině 20 osob bude větší než 105 bodů?
Výsledek: 0,06944.

Výpočet pomocí systému STATISTICA
Uvědomíme si, že výběrový průměr se řídí normálním rozložením se střední hodnotou 100 a rozptylem $\frac{225}{20}$.
Vytvoříme datový soubor o jedné proměnné a o jednom případu. Do Dlouhého jména této proměnné napíšeme
=1 – INormal(105;100;sqrt(225/20)). Zjistíme, že hledaná pravděpodobnost je 0,068.

Příklad 3.: Při provádění určitého pokusu bylo zapotřebí udržovat v laboratoři konstantní teplotu 26,5°C. Teplota byla v jednom pracovním týdnu 46x namátkově kontrolována v různých denních a nočních hodinách. Z výsledků měření byly vypočteny realizace výběrového průměru a výběrové směrodatné odchylky: m = 26,33°C, s = 0,748°C. Za předpokladu, že výsledky měření teploty se řídí rozložením $N(\mu,\sigma^2)$, vypočtěte 95% empirický interval spolehlivosti pro střední hodnotu μ i pro směrodatnou odchylku σ.
Výsledky:
26,11°C < μ < 26,55°C s pravděpodobností aspoň 0,95.
0,62°C < σ < 0,94°C s pravděpodobností aspoň 0,95.

Příklad 4.: U 25 náhodně vybraných dvoulitrových lahví s nealkoholickým nápojem byl zjištěn přesný objem nápoje. Výběrový průměr činil m = 1,99 l a výběrová směrodatná odchylka s = 0,1 l. Předpokládejme, že objem nápoje v láhvi je náhodná veličina s normálním rozložením.
a) Na hladině významnosti 0,05 ověřte tvrzení výrobce, že zákazník není znevýhodněn.
b) Na hladině významnosti 0,05 ověřte tvrzení výrobce, že směrodatná odchylka je 0,08 l.
Výsledky:

Ad a) Řešení pomocí intervalu spolehlivosti: Protože číslo c = 2 leží v intervalu $\left(-\infty, 2{,}0242\right)$ nezamítáme nulovou hypotézu na hladině významnosti 0,05.
Řešení pomocí kritického oboru:
Jelikož hodnota testového kritéria -0,5 neleží v kritickém oboru $\left(-\infty; -1{,}7109\right\rangle$, nezamítáme nulovou hypotézu na hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA
Statistiky – Základní statistiky/tabulky - Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma průměry (normální rozdělení) – zaškrtneme Výběrový průměr vs. Střední hodnota a Jednostr. – do políčka Pr1 napíšeme 1,99 do políčka SmOd1 napíšeme 0,1, do políčka N1 napíšeme 25, do políčka Pr2 napíšeme 2 - Výpočet. Dostaneme p-hodnotu 0,3108, tedy nezamítáme nulovou hypotézu na hladině významnosti 0,05.

ad b) Řešení pomocí intervalu spolehlivosti:
Protože číslo c = 0,08 leží v intervalu $\left(0{,}0781, 0{,}1391\right)$ nezamítáme nulovou hypotézu na hladině významnosti 0,05.
Řešení pomocí kritického oboru:
Jelikož hodnota testového kritéria 37,5 neleží v kritickém oboru $\left(0; 12{,}401\right\rangle \cup \left\langle 39{,}364; \infty\right)$, nejsme oprávněni na hladině významnosti 0,05 zamítnout tvrzení výrobce.

Příklad 5.: Bylo vybráno šest nových vozů téže značky a po určité době bylo zjištěno, o kolik mm se sjely jejich levé a pravé přední pneumatiky. Výsledky: (1,8; 1,5), (1,0; 1,1), (2,2; 2,0), (0,9; 1,1), (1,5; 1,4), (1,6; 1,4). Za předpokladu, že rozdíly uvedených dvojic tvoří náhodný výběr z normálního rozložení s vektorem středních hodnot, testujte na hladině významnosti 0,05 hypotézu, že obě pneumatiky se sjíždějí stejně rychle.
Výsledky:
Řešení pomocí intervalu spolehlivosti: Protože číslo c = 0 leží v intervalu $\left(-0{,}1203, 0{,}287\right)$, nezamítáme nulovou hypotézu na hladině významnosti 0,05.
Řešení pomocí kritického oboru:
Jelikož hodnota testového kritéria 1,0518 neleží v kritickém oboru $\left(-\infty; -2{,}5706\right) \cup \left(2{,}5706; \infty\right)$, nelze na hladině významnosti 0,05 zamítnout nulovou hypotézu.

Výpočet pomocí systému STATISTICA
Vytvoříme nový datový soubor se dvěma proměnnými X a Y a šesti případy. Do proměnných X a Y napíšeme zjištěné hodnoty.
Statistiky – Základní statistiky/tabulky - t-test, závislé vzorky. Zadáme názvy obou proměnných a ve výstupu se podíváme na hodnotu testového kritéria a na p-hodnotu.

Proměnná	t-test pro závislé vzorky (Tabulka1) Označ. rozdíly jsou významné na hlad. p < ,05000							
	Průměr	Sm.odch.	N	Rozdíl	Sm.odch. rozdílu	t	sv	p
X	1,500000	0,489898						
Y	1,416667	0,331160	6	0,083333	0,194079	1,051758	5	0,341062

Protože p-hodnota 0,341062 > 0,05, nezamítáme na hladině významnosti 0,05 hypotézu, že obě přední pneumatiky se sjíždějí stejně rychle.

Příklad 6.: Může politická strana, pro niž se v předvolebním průzkumu vyslovilo 60 z 1000 dotázaných osob, očekávat se spolehlivostí aspoň 0,95, že by v této době ve volbách překročila 5% hranici pro vstup do parlamentu?

Výsledek:

S pravděpodobností přibližně 0,95 platí, že ϑ > 0,047647. Protože tento interval zahrnuje i hodnoty nižší než 0,05, nelze vyloučit, že strana získá méně než 5% hlasů.

Výpočet pomocí systému STATISTICA

Statistiky – Analýza síly testu – Odhad intervalu – Jeden podíl, Z, Chí-kvadrát test – OK – Pozorovaný podíl p: 0,06, Velik. Vzorku (N): 1000, Spolehlivost: 0,9 – Vypočítat. Dostaneme 0,0476.

Příklad 7.: Určitá cestovní kancelář organizuje zahraniční zájezdy podle individuálních přání zákazníků. Z několika minulých let ví, že 30 % všech takto organizovaných zájezdů má za cíl zemi X. Po zhoršení politických podmínek v této zemi se cestovní kancelář obává, že se zájem o tuto zemi mezi zákazníky sníží. Ze 150 náhodně vybraných zákazníků v tomto roce má 38 za cíl právě zemi X. Potvrzují nejnovější data pokles zájmu o tuto zemi? Volte hladinu významnosti 0,05.

Výsledek:

Testové kritérium se realizuje hodnotou -1,24722, která nepatří do kritického oboru, $\left(-\infty, -1{,}645\right\rangle$. H_0 nezamítáme na asymptotické hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA

Vytvoříme datový soubor o dvou proměnných (nazveme je t0 a kvantil) a jednom případu.
Vypočteme realizaci testového kritéria tak, že do Dlouhého jména proměnné t0 napíšeme
=(38/150-0,3)/sqrt(0,3*0,7/150)
Do Dlouhého jména proměnné kvantil napíšeme
=VNormal(0,95;0;1)
Tím získáme kvantil $u_{0,95}$.

	1 t0	2 kvantil
1	-1,24721913	1,644854

Jelikož realizace testového kritéria t_0 = -1,24721913 nepatří do kritického oboru $W = \left(-\infty, -1{,}644854\right\rangle$, H_0 nezamítáme na asymptotické hladině významnosti 0,05.

Příklad 8.: Při výstupní kontrole bylo náhodně vybráno 150 výrobků vyrobených na ranní směně a rovněž 150 výrobků vyrobených na odpolední směně. U ranní směny bylo zjištěno 16 zmetků a u odpolední 12 zmetků. Na asymptotické hladině významnosti 0,05 testujte hypotézu, že pravděpodobnost vyrobení zmetků v obou směnách je táž. Test proveďte pomocí intervalu spolehlivosti a pomocí p-hodnoty.

Výsledek:

S pravděpodobností přibližně 0,95 platí, že –0,039 < $\vartheta_1 - \vartheta_2$ < 0,092. Protože tento interval obsahuje 0, nezamítáme H_0 na hladině významnosti 0,05.

Test pomocí p-hodnoty provedeme ve STATISTICE:

Nejprve vypočteme relativní četnosti vyrobení zmetku na ranní a odpolední směně:

$m_1 = \frac{16}{150} = 0{,}1067$, $m_2 = \frac{12}{150} = 0{,}08$

Statistiky – Základní statistiky a tabulky – Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma poměry – do políčka P 1 napíšeme 0,1067, do políčka N1 napíšeme 150, do políčka P 2 napíšeme 0,08, do políčka N2 napíšeme 150 –Výpočet. Dostaneme p-hodnotu 0,4274, tedy nezamítáme nulovou hypotézu na hladině významnosti 0,05.