Cvičení 5 – příklady u tabule

Přehled vzorců pro náhodný výběr z normálního rozložení
Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozložení $N(\mu, \sigma^2)$. Pak platí

$$M \sim N\left(\mu, \frac{\sigma^2}{n}\right), \text{ tedy } U = \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1),\ K = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),\ T = \frac{M-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1).$$

Interval spolehlivosti pro μ při neznámém σ: $\left(m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1),\ m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) \right)$.

Interval spolehlivosti pro σ při neznámém μ: $\left(\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}} \right)$.

Jednovýběrový t-test: $t_0 = \frac{m-c}{\frac{s}{\sqrt{n}}}$, $W = \left(-\infty, -t_{1-\alpha/2}(n-1)\right\rangle \cup \left\langle t_{1-\alpha/2}(n-1), \infty\right)$.

Test o rozptylu: $t_0 = \frac{(n-1)s^2}{c}$, $W = \left\langle 0, \chi^2_{\alpha/2}(n-1)\right\rangle \cup \left\langle \chi^2_{1-\alpha/2}(n-1), \infty\right)$.

Párový t-test: Nechť $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \ldots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ je náhodný výběr z dvourozměrného rozložení, $n \geq 2$.

Označíme $\mu = \mu_1 - \mu_2$ a zavedeme rozdílový náhodný výběr $Z_1 = X_1 - Y_1, \ldots, Z_n = X_n - Y_n$, o němž předpokládáme, že pochází z normálního rozložení. Test hypotézy o rozdílu středních hodnot $\mu_1 - \mu_2$ se nazývá párový t-test a provádí se stejně jako jednovýběrový t-test aplikovaný na rozdílový náhodný výběr

Příklad 1.: (viz př. 6.4.1. ze skript) 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?

Příklad 2.: (viz př. 6.4.4. ze skript) 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.

Příklad 3.: (viz př. 6.4.6. ze skript) 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. Na hladině významnosti 0,05 ověřte tvrzení výrobce, že směrodatná odchylka je 0,08 l.

Příklad 4.: (viz př. 6.4.7. ze skript) 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.

Přehled vzorců pro dva nezávislé náhodné výběry z normálních rozložení

Interval spolehlivosti pro podíl rozptylů: $\left(\frac{s_1^2/s_2^2}{F_{1-\alpha/2}(n_1-1,n_2-1)},\frac{s_1^2/s_2^2}{F_{\alpha/2}(n_1-1,n_2-1)}\right)$.

Interval spolehlivosti pro rozdíl středních hodnot při shodných rozptylech:

$(m_1 - m_2 - s_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\, t_{1-\alpha/2}(n_1+n_2-2),\ m_1 - m_2 + s_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\, t_{1-\alpha/2}(n_1+n_2-2))$, kde

$$s_*^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}.$$

Příklad 5.: (viz př. 7.3.1. ze skript) Bylo vylosováno 11 stejně starých selat téhož plemene. Šesti z nich byla předepsána výkrmná dieta č. 1 a zbylým pěti výkrmná dieta č. 2. Průměrné denní přírůstky v Dg za dobu půl roku jsou následující:

dieta č. 1: 62, 54, 55, 60, 53, 58
dieta č. 2: 52, 56, 49, 50, 51.

Zjištěné hodnoty považujeme za realizace dvou nezávislých náhodných výběrů pocházejících z rozložení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$.
a) Sestrojte 95% empirický interval spolehlivosti pro podíl rozptylů.
b) Za předpokladu, že data pocházejí z rozložení $N(\mu_1, \sigma^2)$ a $N(\mu_2, \sigma^2)$, sestrojte 95% empirický interval spolehlivosti pro rozdíl středních hodnot $\mu_1 - \mu_2$.
Pro usnadnění výpočtů máte k dispozici následující číselné charakteristiky:
$m_1 = 57$, $m_2 = 51{,}6$, $s_1^2 = 12{,}8$, $s_2^2 = 7{,}3$.