

# Statistika (STT)

Zápočtová práce komplet 28. 02. 2020

Zadání č. 1

Jméno a Příjmení: \_\_\_\_\_

Číslo studenta: \_\_\_\_\_

Podpis: \_\_\_\_\_ (stvrzuji, že jsem práci vypracoval samostatně)

1. (5 b.) Byla změřena výška (v cm) 77 studentů prvního ročníku univerzity, viz níže. Spočítejte jejich a) průměrnou výšku, dále b) medián, c) modus, d) 70%-ní kvantil, e) mezikvartilové rozpětí a f) směrodatnou odchylku.

176 168 169 184 194 178 177 173 176 171 179 182 171 173 169 166 177 172 175  
180 179 173 180 170 179 176 176 174 164 170 180 190 177 162 201 183 178 182  
173 176 176 181 188 161 172 168 169 161 163 175 177 180 175 194 176 167 193  
160 184 190 179 163 159 179 185 173 169 178 180 171 170 161 167 170 178 181  
168

2. (5 b.) Skupina studentů byla podrobena testu. Kolik jednotliví studenti obdrželi bodů shrnuje následující tabulka. Spočítejte a) průměrný bodový zisk, dále b) medián, c) modus, d) 80%-ní kvantil, e) mezikvartilové rozpětí a f) směrodatnou odchylku.

počet bodů	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
počet studentů	12	25	105	198	71	64	68	49	37	24	12	7	5	2	1

3. (5 b.) Hodnota určité veličiny byla změřena na objektech skupiny A a skupiny B. V obou skupinách zvlášť spočítejte a) průměr hodnot, b) medián, c) směrodatnou odchylku, d) mezikvartilové rozpětí a e) nakreslete boxplot. f) Na základě těchto údajů obě skupiny hodnot porovnejte (tj. uveďte zda je některá skupina zjevně vyšší nebo zda jsou srovnatelné) z hlediska jejich polohy a variability.

SkupinaA: 16.6 19.5 19.5 24.6 20.4 20.7 24.3 17.6 24.2 21.6 13.5 17.5 18.2  
19.1 12.8 18.6 17.7 20.8 16.5 19.2 12.4 25.4 20.4 16.1 22.2 18.9 18.9 18.7  
22.1 24.9 19.4 14.1 19.1 16.1

SkupinaB: 12.5 5.0 8.2 0.9 6.7 5.4 3.7 6.7 15.3 11.2 6.6 11.2 7.1  
12.5 3.4 5.2 -3.2 0.6 10.5 -0.6 12.1 6.9 8.9 4.1 5.8 6.3 7.1 6.8  
6.5 7.1 8.7

4. (5 b.) Následující data se týkají skupiny uchazečů o práci u policie. U každého uchazeče byla zjištěna jeho výška (v cm) a reakční doba (v sekundách).  
a) určete korelační koeficient mezi výškou a reakční dobou těchto uchazečů  
b) určete parametry lineární závislosti reakční doby na výšce  
c) pomocí modelu z části b) odhadněte reakční dobu uchazeče vysokého 174 cm.

uchazeč č.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
výška	171.9	171.3	172.5	191.9	181.8	171.1	171.7	185.2	176.3	177.0	161.4	169.3	186.2	181.0	176.9
reakční doba	0.304	0.273	0.343	0.261	0.316	0.302	0.275	0.255	0.271	0.260	0.283	0.274	0.344	0.288	0.257

5. (5 b.) V každém kole hazardní hry s pravděpodobností 0.41 vyhraje 10 Kč (zisk 10 Kč) a v opačném případě ztratíte 10 Kč (zisk -10 Kč).  
a) Doplňte tabulku rozdělení pravděpodobnosti vašeho celkového zisku  $X$  po 3 kolech.

$x_i$	-30	-20	-10	0	10	20	30
$P(X = x_i)$							

- b) Spočítejte pravděpodobnost, že po 3 kolech budete v zisku.  
c) Spočítejte výši středního (očekávaného) celkového zisku.  
d) Spočítejte směrodatnou odchylku celkového zisku.
6. (5 b.) Experimentálně bylo zjištěno, že okolní teplota (ve  $^{\circ}\text{C}$ ), při které dojde k nenávratnému zničení harddisku, je spojitá náhodná veličina s hustotou:

$$f(x) = \begin{cases} c \cdot (x - 89) & \text{pro } x \in [89, 97] \\ 0 & \text{jinak.} \end{cases}$$

- a) Určete konstantu  $c$ .  
b) Spočítejte střední hodnotu této náhodné veličiny.  
c) Spočítejte rozptyl této náhodné veličiny.  
d) Víme, že okolní teplota vystoupí až na  $93^{\circ}\text{C}$ . S jakou pravděpodobností dojde ke zničení harddisku?  
e) Na jakou hodnotu maximálně může teplota vystoupit, aby pravděpodobnost zničení byla nejvýše 40%?
7. (5 b.) V urně jsou balonky třech různých barev: 7 bílých, 1 černý a 6 zelených. Náhodně jste vytáhli s vrácením 7 balonků.  
a) S jakou pravděpodobností bylo mezi vytaženými právě 6 bílých?  
b) S jakou pravděpodobností byly mezi vytaženými alespoň 3 bílé?  
c) Určete střední (očekávaný) počet bílých balonků mezi vytaženými.
8. (5 b.) V urně jsou balonky třech různých barev: 6 bílých, 1 černý a 5 zelených. Náhodně vytahujete s vrácením balonky až do té doby, kdy poprvé vytáhnete bílý balonek.  
a) S jakou pravděpodobností budete muset táhnout právě 7-krát (bílý bude vytažen poprvé v 7. tahu)?  
b) S jakou pravděpodobností budete muset táhnout více než 3-krát?  
c) Určete střední (očekávaný) počet "neúspěšných" tahů před prvním vytaženým bílým balonkem.
9. (5 b.) V urně jsou balonky třech různých barev: 8 bílých, 5 černých a 4 zelené. Náhodně jste vytáhli bez vrácení 9 balonků.  
a) S jakou pravděpodobností bylo mezi vytaženými právě 7 bílých?  
b) S jakou pravděpodobností bylo mezi vytaženými alespoň 7 bílých?  
c) Určete střední (očekávaný) počet bílých balonků mezi vytaženými.
10. (5 b.) Zajímají nás příchody zákazníků do obchodu. Je odpozorováno, že v průměru přijde do obchodu 5 zákazníků za hodinu, tj. v průměru se čeká na dalšího zákazníka 12 minut.  
a) S jakou pravděpodobností během 40 minut přijde do obchodu právě 7 zákazníků?  
b) S jakou pravděpodobností během 40 minut přijde do obchodu alespoň 5 zákazníků?
11. (5 b.) Zajímají nás příchody zákazníků do obchodu. Je odpozorováno, že v průměru přijde do obchodu 6 zákazníků za hodinu, tj. v průměru se čeká na dalšího zákazníka 10 minut.

- a) S jakou pravděpodobností bude doba čekání na příchod dalšího zákazníka delší než 30 minut?
- b) S jakou pravděpodobností bude doba čekání na příchod dalšího zákazníka ležet v intervalu (20, 57) minut?
12. (5 b.) K měření hmotnosti máte k dispozici zkalibrovanou váhu. Víte, že naměřená hodnota je náhodná veličina s normálním rozdělením o střední hodnotě rovné hmotnosti váženého tělesa a směrodatnou odchylkou rovnou 1.7 g. Jestliže je skutečná hmotnost váženého tělesa 37.4 g, určete následující:
- a) Jaká je pravděpodobnost, že naměřená hodnota bude větší než 35.9 g?
- b) Jaká je pravděpodobnost, že naměřená hodnota bude menší než 37.5 g?
- c) Jaká je pravděpodobnost, že naměřená hodnota bude ležet v intervalu (34.7, 37.1) g?
- d) Nad jakou hodnotou bude naměřená hmotnost s pravděpodobností 0.75?
- e) Pod jakou hodnotou bude naměřená hmotnost s pravděpodobností 0.55?
13. (5 b.) Máte k dispozici opakovaná měření tělesné teploty (ve °C):
- 38.5 38.6 38.6 38.5 38.6 38.7 38.3 38.4 38.5 38.3 38.5 38.4 38.5 38.4 38.6  
38.5 38.5 38.6 38.6 38.6 38.4 38.5 38.5 38.6 38.8 38.5
- Víte, že rozptyl měřícího přístroje je roven  $0.01 \text{ } ^\circ\text{C}^2$ .
- a) Vypočtete 99%-ní interval spolehlivosti pro tělesnou teplotu a výsledek interpretujte.
- b) Doplňte předpoklady použitého postupu.
14. (5 b.) Máte k dispozici opakovaná měření tělesné teploty (ve °C):
- 39.7 40.4 39.5 39.7 39.9 40.0 39.8 39.7 40.0 39.6 39.7 39.6 39.9 39.8 39.8  
39.7 39.7 39.7 39.7 39.9 39.7 40.0 39.7 39.7 39.9 39.7 39.7 39.6 39.9 39.6  
40.1
- a) Vypočtete 90%-ní interval spolehlivosti pro tělesnou teplotu a výsledek interpretujte.
- b) Doplňte předpoklady použitého postupu.
15. (5 b.) Převážení několika balíčků mouky byly zjištěny následující odchylky (v gramech) od přesné hodnoty 1 kg:
- 1.0 1.1 -3.4 1.2 -2.3 1.3 -0.7 1.0 0.8 0.3 -1.8 1.3 -1.7 0.2 3.7  
2.4 -0.8 -0.7 3.0 -2.6 0.4 0.2 0.3 -1.8 0.1 0.8 2.2 1.8 0.1 -1.0  
-1.6 -1.7 1.8 1.5 -1.9 2.8 0.4 -2.4 0.1 0.8 0.8 -0.2 -0.3
- a) Určete 95%-ní interval spolehlivosti pro rozptyl hmotnosti v jednom balíčku a výsledek interpretujte.
- b) Doplňte předpoklady použitého postupu.
16. (5 b.) Převážení několika balíčků mouky byly zjištěny následující odchylky (v gramech) od přesné hodnoty 1 kg:
- 2.0 1.9 -1.3 4.0 -0.5 -1.1 0.7 6.1 2.9 0.7 2.3 1.4 1.6 3.8 -0.1  
3.0 -0.3 0.9 1.3 -0.9 -0.3 0.5 1.5 3.0 0.0 0.4 -0.3 -0.7 -0.9 2.7  
-1.7 -0.1 -2.0 2.8 -0.4 1.0 1.1 2.8 0.2 0.9 1.0 0.4 -0.9 4.1 1.2  
-2.4 -1.2 -0.9 1.1 -0.7
- a) Určete 99%-ní interval spolehlivosti pro pravděpodobnost, že náhodně vybraný balíček bude lehčí než 1 000 g a výsledek interpretujte.
- b) Doplňte předpoklady použitého postupu.
17. (5 b.) Výrobní zařízení produkuje šrouby o délce, jejíž rozptyl je roven  $0.1 \text{ mm}^2$ . Zařízení by mělo produkovat šrouby o střední délce 3 cm. Přeměření délky několika výrobků vedlo k těmto hodnotám (v cm):

2.97 3.08 3.02 3.01 3.05 3.03 3.04 2.99 3.04 3.05 3.02 3.00 3.04 3.00 2.98  
3.01 3.03 3.08 3.01 3.03

- a) Na hladině významnosti  $\alpha = 0.05$  rozhodněte, zda je potřeba zařízení nově seřadit.  
b) Doplňte předpoklady použitého postupu.

18. **(5 b.)** Výrobní zařízení produkuje šrouby, jejíž střední délka by měla být rovna 2.5 cm. Přeměření délky několika výrobků vedlo k těmto hodnotám (v cm):

2.44 2.47 2.51 2.49 2.52 2.50 2.49 2.52 2.47 2.50 2.55 2.49 2.46 2.50 2.51  
2.55 2.51 2.53 2.50 2.52 2.51 2.50

- a) Na hladině významnosti  $\alpha = 0.1$  rozhodněte, zda je potřeba zařízení nově seřadit.  
b) Doplňte předpoklady použitého postupu.

19. **(5 b.)** Na 26 pacientech byl zkoumán vliv dvou léků na krevní tlak. Pacienti byli náhodně rozděleni do dvou skupin. Pacienti v první skupině dostali lék A a v druhé skupině lék B. Data v tabulce udávají systolický tlak (v mm Hg) u pacientů v určitém čase po podání léku:

lék A	158	175	169	150	149	136	195	178	182	200	196	185	153
lék B	172	173	217	188	189	172	195	175	236	204	199	215	197

- a) Na hladině významnosti  $\alpha = 0.01$  rozhodněte, zda existuje rozdíl mezi účinky obou léků na krevní tlak.  
b) Doplňte předpoklady použitého postupu.

20. **(5 b.)** Na 11 pacientech byl zkoumán vliv dvou léků na krevní tlak. Všichni pacienti nejprve dostali lék A a po určitém čase lék B. Data v tabulce udávají systolický tlak (v mm Hg) u pacientů po podání léku:

lék A	170	158	163	175	182	169	188	155	154	161	158
lék B	206	183	207	187	194	210	225	189	199	194	196

- a) Na hladině významnosti  $\alpha = 0.01$  rozhodněte, zda existuje rozdíl mezi účinky obou léků na krevní tlak.  
b) Doplňte předpoklady použitého postupu.

21. **(5 b.)** Uvádí se, že každá pátá výhra v loterii je vyšší než 6800 Kč. Ověřte tento údaj na hladině významnosti  $\alpha = 0.1$ , jestliže víte, že poslední výhry v loterii byly rovny (v Kč):

1100 1000 800 2800 1900 2000 200 1800 100 200 600 1500  
1000 3200 500 700 200 2200 400 900 9500 1900 1700 2800  
200 1000 600 3200 2700 300 1700 4900 400 1300 600 2200  
22300 4500 2400 900 400 3500 5800 2600 400 400 2700 900  
600 500 500 1300 500 300 500 200 1000 700 600 73600  
4000 600 1400 300 700 700 800 2500 200 2200 300 800  
1400 17500

Výsledek slovně interpretujte.

22. **(5 b.)** Trenér o své svěřenkyňi, diskařce, tvrdí, že "v průměru" hodí 50 metrů. Ověřte jeho tvrzení na hladině významnosti  $\alpha = 0.1$ , jestliže víte, že poslední hody diskařky v závodech byly rovny (v m):

34.16 61.73 51.98 56.86 49.63 62.10 61.71 51.57 52.11 7.77 41.16 50.39  
53.20 61.60 56.20 0.00 59.95 54.45 38.14 46.58 59.22 50.94 52.27 56.00  
52.60 62.82 48.38 48.42 58.27 51.10 46.79 64.55 56.87 56.42 49.28 59.80  
51.89 43.69 49.58 38.29 62.21 56.34 38.02 52.90 56.43

Zdůvodněte, proč jste zvolili použitý postup a výsledek slovně interpretujte.

23. **(5 b.)** Ze vzorků typu A bylo kvalitních (vydrželo zátěžový test) 79 a bylo nekvalitních (nevydrželo zátěžový test) 41. Ze vzorků typu B stejný test vydrželo 141 a 131 test nevydrželo. Na hladině významnosti 0.05 rozhodněte, zda existuje souvislost mezi typem vzorku a jeho kvalitou.
24. **(5 b.)** Při různých teplotách (ve °C) byl naměřen následující obsah křemíku v surovém železe (v %).

teplota	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510
obsah	0.44	0.37	0.62	0.50	0.48	0.53	0.66	0.81	0.68	0.62	0.66	0.86

- a) Odhadněte lineární regresní přímku závislosti obsahu křemíku na teplotě.
- b) Jaká část celkové proměnlivosti obsahu křemíku je vysvětlena lineární závislostí na teplotě?
- c) Za předpokladu, že data nepocházejí z normálního rozdělení, na hladině významnosti  $\alpha = 0.05$  rozhodněte, zda obsah křemíku v surovém železe významně závisí na teplotě.