Vaje

1. Računanje standardnega odmika ter intervala zaupanja

S spektrofotometrično metodo določimo naslednje vsebnosti železa v aluminiju: 23.3, 24.1, 23.6, 23.8, 23.5, 23.6, 23.5 in 23.7 ppm. Izračunajte standardni odmik in interval zaupanja pri 5% tveganju!

Najprej izračunamo poprečno vrednost ($\bar{x}$) in nato kvadrate razlik med poprečjem in posamezno meritvijo.

$$\bar{x} = \frac{23.3 + 24.1 + 23.6 + 23.8 + 23.5 + 23.6 + 23.7 + 23.5}{8} = 23.6375$$

Tabela 1. Računanje standardnega odmika

x_i	$(x_i - \bar{x})^2$
23.3	0.113906
24.1	0.213906
23.6	0.001406
23.8	0.026406
23.5	0.018906
23.6	0.001406
23.7	0.003906
23.5	0.018906
Vsota	0.39875

Vsoto delimo s številom prostotnih stopenj (n-1 = 7) in kvocient nato korenimo.

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{0.39875}{7}} = 0.239$$

Standardni odmik je 0.239.

Pri drugem načinu računanja standardnega odmika najprej izračunamo vsoto kvadratov ter kvadrat vsote meritev.

$$\sum_{i=1}^{n} x_i^2 = 23.3^2 + 24.1^2 + 23.6^2 + 23.8^2 + 23.5^2 + 23.6^2 + 23.7^2 + 23.5^2 = 4470.25$$

$$(\sum_{i=1}^{n} x_i)^2 = (23.3 + 24.1 + 23.6 + 23.8 + 23.5 + 23.6 + 23.7 + 23.5)^2 = 35758.81$$

Nato izračunamo standardni omik po naslednji formuli:

$$s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2}{n-1}} = \sqrt{\frac{4470.25 - \frac{1}{8}35758.81}{7}} = 0.239$$

Ko imamo izračunano poprečno vrednost in standardni odmik napišimo še interval zaupanja pri 5% tveganju.

Spodnja meja: $\bar{x} - \frac{z \cdot \sigma}{\sqrt{n}} = 23.6 - \frac{1.96 \cdot 0.239}{\sqrt{8}} = 23.4$

Zgornja meja: $\bar{x} + \frac{z \cdot \sigma}{\sqrt{n}} = 23.6 + \frac{1.96 \cdot 0.239}{\sqrt{8}} = 23.8$

2. Primerjava dveh velikih serij meritev; n_1 in $n_2 \geq 30$

V medlaboratorijski primerjalni študiji določamo vsebnost železa v biološkem referenčnem materialu. Prvi laboratorij izmeri poprečno vrednost železa v vzorcu 32.6

ppm in s_1^2 =6.55 (n_1=36). Rezultati drugega laboratorija so $\overline{x_2} = 31.6$, s_2^2 =4.05 (n_2=30). Ali so rezultati analiz iz obeh laboratorijev enaki?

Preden se lotimo primerjave dveh popreǎnih vrednosti moramo narediti obojestranski F-test pri α=0.05, da preverimo, če varianci pripadata isti populaciji. V števec postavimo večjo varianco. Ničelna hipoteza se glasi: varianci se signifikantno ne razlikujeta.

$$F = \frac{s_1^2}{s_2^2} = \frac{6.55}{4.05} = 1.62$$

V tabelah poiščemo kritično vrednost, ki je $F_{0.05,35,29} = 2.056$*. Ker je kritična vrednost večja kot pa izračunana obdržimo ničelno hipotezo in se lotimo primerjave poprečnih vrednosti.*
Ker je število meritev v obeh serijah večje ali enako 30 in ker se varianci med seboj bistveno ne razlikujeta, uporabimo za primerjavo poprečnih vrednosti z-test.

$$z = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\left|32.6 - 31.6\right|}{\sqrt{\frac{6.55}{36} + \frac{4.05}{30}}} = \frac{1}{0.563} = 1.78$$

Ker je izračunana vrednost manjša od kritične $z_{0.05}$*=1.96, obdržimo ničelno hipotezo, da ni bistvenih razlik med poprečnima vrednostima pri 5% tveganju.*

3. Primerjava dveh majhnih serij meritev; n_1 in n_2 < 30

a) Laboratorij določa vsebnost nečistoč v aluminiju. V dveh vzorcih določa prisotnost železa. Rezultati so podani v tabeli (μg/g). Ali sta vzorca enaka?

Vzorec 1	Vzorec 2
3.46	3.35

3.56	3.42
3.48	3.33
3.52	3.45
3.51	3.38

Ponovno najprej z F-testom preverimo ali sta varianci enaki, pred tem pa izračunamo poprečne vrednosti in standardne odmike.

$$\overline{x}_1 = \frac{3.46 + 3.56 + 3.48 + 3.52 + 3.51}{5} = \frac{17.53}{5} = 3.51$$

$$\overline{x}_2 = \frac{3.35 + 3.42 + 3.33 + 3.45 + 3.38}{5} = \frac{16.93}{5} = 3.39$$

$$s_1 = \sqrt{\frac{(3.46 - 3.51)^2 + (3.56 - 3.51)^2 + (3.48 - 3.51)^2 + (3.52 - 3.51)^2 + (3.51 - 3.51)^2}{4}} = 0.0385$$

$$s_2 = \sqrt{\frac{(3.35 - 3.39)^2 + (3.42 - 3.39)^2 + (3.33 - 3.39)^2 + (3.45 - 3.39)^2 + (3.38 - 3.39)^2 +}{4}} = 0.0493$$

$$F = \frac{0.0493^2}{0.0385^2} = \frac{0.00243}{0.00148} = 1.64$$

Kritična vrednost za F(0.05,4,4)=9,6. Ker je izračunana vrednost manjša od kritične, obdržimo ničelno hipotezo, da ni bistvenih razlik med variancama. Sedaj se lahko lotimo primerjave poprečnih vrednosti. Ker je število meritev manjše od 30 uporabimo za primerjavo poprečnih vrednosti t-test.

Najprej izračunamo skupni standardni odmik

$$s = \sqrt{\frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{4\cdot 0.00148 + 4\cdot 0.00243}{5 + 5 - 2}} = 0.0442$$

in nato t vrednost.

$$t=\frac{\left|\overline{x}_1-\overline{x}_2\right|}{\sqrt{s^2\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}=\frac{\left|3.51-3.39\right|}{\sqrt{0.0442^2\cdot\left[\frac{1}{5}+\frac{1}{5}\right]}}=\frac{0.12}{0.028}=4.29$$

Kritična vrednost za t pri 5% tveganju in 8 prostostnih stopnjah je 2.306. Ker je kritična vrednost manjša od izračunane se ničelna hipoteza zavrže in privzame alternativna hipoteza, da obstajajo bistvene razlike med vzorcema.

b) Forenzični laboratorij mora določiti vsebnost cinka v kontrolni barvi ter v barvi s kraja prometne nesreče. V kontrolni barvi določijo naslednje vsebnosti: 3.2, 3.4,3.0, 2.8 in 3,5 ppm. V laku s kraja prometne nesreče pa določijo 3.1, 3.15, 3.05 in 3.1 ppm. Ali gre za identičen vzorec?

Najprej izračunamo povprečne vrednosti, standardne odklone ter F-test.

$$\overline{x_1}=\frac{3.2+3.4+3.0+2.8+3.5}{5}=\frac{15.9}{5}=3.18$$

$$\overline{x}_2=\frac{3.1+3.15+3.05+3.1}{4}=\frac{12.4}{4}=3.1$$

$$s_1=\sqrt{\frac{(3.2-3.18)^2+(3.4-3.18)^2+(3.0-3.18)^2+(2.8-3.18)^2+(3.5-3.18)^2}{4}}=0.286$$

$$s_2 = \sqrt{\frac{(3.2-3.1)^2 + (3.4-3.1)^2 + (3.0-3.1)^2 + (2.8-3.1)^2}{3}} = 0.041$$

$$F = \frac{s_1^2}{s_2^2} = \frac{0.082}{0.001667} = 49.2$$

Kritična vrednost $F_{0.05,4,3}$=15.1 je manjša od izračunane, zato zavržemo ničelno hipotezo in predpostavimo, da varianci nista enaki. Ker pogoj enakih varianc ni izpolnjen, t-testa ne moremo izvesti.

4. Določevanje potrebnega števila meritev

a) Določite število meritev vsebnosti vitamina C v prehrambenem izdelku, če koncentracija omenjenega vitamina ne sme odstopati od deklarirane vrednosti za več kot 3 μg/g pri 95% zaupanju. Standardni odmik metode je 11.5 μg/g.

Iz tabele odčitamo z vrednost pri 5% tveganju in znaša 1.96. Prav tako si napišimo enačbo za interval zaupanja.

$$\bar{x} = \mu \pm \frac{z \cdot \sigma}{\sqrt{n}} \Rightarrow n = \left(\frac{z \cdot \sigma}{\bar{x} - \mu} \right)^2 = \left(\frac{11.5 \cdot 1.96}{3} \right)^2 = 56.5$$

Opravimo 57 meritev.

b) Imamo enako nalogo, vendar uporabimo analizno metodo pri kateri je nižji standardni odmik in znaša 4.5.

Začetek je enak kot pri prejšnjem primeru.

$$\bar{x} = \mu \pm \frac{z \cdot \sigma}{\sqrt{n}} \Rightarrow n = \left(\frac{z \cdot \sigma}{\bar{x} - \mu}\right)^2 = \left(\frac{4.5 \cdot 1.96}{3}\right)^2 = 8.6$$

Opraviti bi morali 9 meritev. V tem primeru ne moremo uporabiti normalne porazdelitve zato si odčitamo iz tabele t vrednost pri 8 prostostnih stopnjah. Kritična t vrednost je 2.306. Nato ponovimo postopek.

$$n = \left(\frac{z \cdot \sigma}{\bar{x} - \mu}\right)^2 = \left(\frac{4.5 \cdot 2.306}{3}\right)^2 = 11.96$$

Sedaj iteriramo postopek dokler se nam n spreminja iz ene do druge ponovitve.

$t_{0.05,11} = 2.201$

$$n = \left(\frac{z \cdot \sigma}{\bar{x} - \mu}\right)^2 = \left(\frac{4.5 \cdot 2.201}{3}\right)^2 = 10.9$$

$t_{0.05,10} = 2.228$

$$n = \left(\frac{z \cdot \sigma}{\bar{x} - \mu}\right)^2 = \left(\frac{4.5 \cdot 2.228}{3}\right)^2 = 11.2$$

Vidimo, da se n ne spreminja več, zato za število meritev določimo bolj konzervativno vrednost to je 12 paralelk.

5. Določevanje ubežnikov

Naredili smo serijo meritev koncentracij mangana v jeklu s spektrofotometričmo metodo. Dobili smo naslednjo serijo meritev: 0.41, 0.42, 0.42, 0.44 in 0.54%. Določite ali obstaja ubežnik v naših meritvah!

Za določitev ubežne točke moramo naprej izračunati poprečno vrednost in standardni odklon.

$$\bar{x} = \frac{0.41 + 0.42 + 0.42 + 0.44 + 0.54}{6} = \frac{2.26}{6} = 0.446$$

$$s = \sqrt{\frac{(0.41 - 0.446)^2 + (0.42 - 0.446)^2 + (0.42 - 0.446)^2 + (0.44 - 0.446)^2 + (0.54 - 0.446)^2}{4}} =$$

$$= \sqrt{\frac{0.01152}{4}} = 0.0537$$

Nato uporabimo Grubbsov test za izločitev ubežnikov. Za ubežno točko sumimo meritev, ki najbolj odstopa od ostalih paralelk.

$$G = \frac{x_i - \bar{x}}{s} = \frac{0.54 - 0.446}{0.0537} = 1.75$$

Na koncu primerjamo izračunano vrednost s kritično pri 5% tveganju , ki jo dobimo v tabelah. Kritična vrednost znaša 1.481. Ker je izračunana vrednost večja od kritične zavržemo ničelno hipotezo, da so vse meritve iz iste populacijo in sprejmemo alternativno, da je izbrana meritev ubežnik.

6. Primerjava dveh porazdelitev s χ^2 testom

S plinsko kromatografijo smo analizirali vzorec v katerih smo našli 200 organskih spojin. Spojine smo grupirali v 9 enako širokih razredov glede na retencijski čas. Ali so organske spojine enakomerno porazdeljene po celotnem kromatogramu?

F_1^o	F_2^o	f_3^o	f_4^o	f_5^o	f_6^o	f_7^o	f_8^o	f_9^o
20	24	27	18	17	24	14	25	31

Naša hipoteza se glasi, da so spojine enkomerno porazdeljene po kromatogramu, zato je v vsakem segmentu teoretično celotno število spojin deljeno s številom segmentov.

$$f_i^T = \frac{\text{stevilo vseh spojin}}{\text{stevilo segmentov}} = \frac{200}{9} = 22.22$$

Za testiranje ničelne hipoteze uporabimo χ^2 test.

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i^o - f_i^T)^2}{f_i^T} = \frac{(20-22.22)^2}{22.22} + \frac{(24-22.22)^2}{22.22} + \frac{(27-22.22)^2}{22.22} + \frac{(18-22.22)^2}{22.22} +$$

$$+ \frac{(17-22.22)^2}{22.22} + \frac{(24-22.22)^2}{22.22} + \frac{(14-22.22)^2}{22.22} + \frac{(25-22.22)^2}{22.22} + \frac{(31-22.22)^2}{22.22} =$$

$$= 0.222 + 0.142 + 1.027 + 0.802 + 1.227 + 0.142 + 3.042 + 0.347 + 3.467 = 10.42$$

Kritična vrednost pri 8 prostostnih stopnjah in α=0.05 je 17.5. Ker je izračunana vrednost nižja od kritične sledi, da porazdelitev organskih spojin v kromatogramu ni bistveno različna od enakomerne porazdelitve.

7. Kalibracijska premica

V vzorcu hrane določamo vsebnost vitamina C s tekočinsko kromatografijo. Za kvantitativno ovrednotenje rezultatov si pripravimo umeritveno krivuljo. Dobimo naslednje rezultate.

C (μg/l)	površina
0.1	1.78
0.3	2.45
0.5	3.1
0.7	3.45
1.3	5.3
2	7.3

Izračunajte koeficiente premice, napake posameznih parametrov ter interval zaupanja!

Najprej si pripravimo tabelo iz katere bomo izračunali željene parametre.

k	x_j	y_j	$x_j y_j$	x_j^2
1	0.1	1.78	0.178	0.01
2	0.3	2.45	0.735	0.09
3	0.5	3.1	1.55	0.25
4	0.7	3.45	2.415	0.49
5	1.3	5.3	6.89	1.69
6	2	7.3	14.6	4
Vsota	4.9	23.38	26.368	6.53

Iz tako pripravljene tabele nato izračunamo naklon in odsek premice.

$$y = a + bx$$

$$b = \frac{\sum_{j=1}^{k}(x_j - \bar{x})\cdot(y_j - \bar{y})}{\sum_{j=1}^{k}(x_j - \bar{x})^2} = \frac{k\sum_{j=1}^{k} x_j y_j - \sum_{j=1}^{k} x_j \sum_{j=1}^{k} y_j}{k\sum_{j=1}^{k} x_j^2 - (\sum_{j=1}^{k} x_j)^2} = \frac{6\cdot 26.368 - 4.9\cdot 23.38}{6\cdot 6.53 - 4.9^2} = 2.877$$

$$a = \bar{y} - b\bar{x} = \frac{\sum_{j=1}^{k} y_j \sum_{j=1}^{k} x_j^2 - \sum_{j=1}^{k} x_j \sum_{j=1}^{k} x_j y_j}{k\sum_{j=1}^{k} x_j^2 - (\sum_{j=1}^{k} x_j)^2} = \frac{23.38\cdot 6.53 - 4.9\cdot 26.368}{6\cdot 6.53 - 4.9^2} = 1.547$$

Celotna napaka:

Naredimo si tabelo $(y_j - \hat{y})^2$ *vrednosti. Za računanje* $\hat{y}$ *vrednosti uporabimo pravkar izračunane koeficiente premice a in b.*

	Y	$\hat{y}$	$(y_j - \hat{y})^2$	x_i	$(x_j - \bar{x})^2$
1	1.78	1.83	0.00299	0.1	0.5136
2	2.45	2.41	0.00159	0.3	0.2669
3	3.1	2.99	0.01311	0.5	0.1003
4	3.45	3.56	0.01230	0.7	0.0136
5	5.3	5.29	0.00017	1.3	0.2336
6	7.3	7.30	0.00000	2	1.4003

$$s_e = \sqrt{\frac{\sum (y_j - \hat{y})^2}{k-2}} = \sqrt{\frac{0.0316}{4}} = 0.087$$

Napaka odklona in odseka:

$$s_b = \sqrt{\frac{s_e^2}{\sum_{j=1}^{k}(x_j - \bar{x})^2}} = \sqrt{\frac{0.0754}{2.528}} = 0.173$$

$$s_a = s_e \sqrt{\frac{\sum_{j=1}^{k} x_j^2}{n\sum_{j=1}^{k}(x_j - \bar{x})^2}} = s_b \sqrt{\frac{\sum_{j=1}^{k} x_j^2}{n}} = 0.087\sqrt{\frac{6.53}{6 \cdot 2.528}} = 0.057$$

$$y_0 = a + bx_0 \pm t(\alpha, k-2)s_e \sqrt{\frac{1}{k} + \frac{(x_0 - \bar{x})^2}{\sum_{j=1}^{k}(x_j - \bar{x})^2}}$$

Izmerite površine kromatografskih vrhov:3.1, 3.15, 3.2. Določite koncetracijo vitamina C v neznanem vzorcu.

$$x_s = \frac{\bar{y}_s - a}{b} \pm t(\alpha, k-2)\frac{s_e}{b}\sqrt{\frac{1}{n_s} + \frac{1}{k} + \frac{(\bar{y}_s - \bar{y})^2}{b^2 \cdot \sum_{j=1}^{k}(x_j - \bar{x})^2}}$$

8. Grupiranje objektov

V laboratoriju smo določevali vsebnost dveh maščobnih kislin v šestih sirih. Dobimo naslednje vrednosti v υg/g: S_1(2.4,4.5), S_2(2.3,4.0), S_3(1.4,5.5), S_4(6.2,3.5), S_5(5.4,3.5) in S_6(3.4,2.5). Grupirajte podatke! Uporabite Manhattansko razdaljo in najmanjšo razdaljo med grupami.