Uppgift 10.2

Två guldgruvor jämförs för att undersöka hur ekonomin i brytning av guld har utvecklats över tid. Stickprovsmedelvärdet från en gammal gruva är $\bar{x}_1 = 0.233$ ounces guld per ton malm och stickprovsmedelvärdet från en ny gruva är $\bar{x}_2 = 0.127$ ounces guld per ton malm. Skatta skillnaden $\mu_1 - \mu_2$.

Lösning:

Sökt: $\widehat{\mu_1 - \mu_2}$

$\hat{\mu}_1 = \bar{x}_1$, $\hat{\mu}_2 = \bar{x}_2$

Sida 337 : $\widehat{\mu_1 - \mu_2} = \hat{\mu}_1 - \hat{\mu}_2 = \bar{x}_1 - \bar{x}_2 = 0.233 - 0.127 = \boxed{0.106}$

Uppgift 10.8

En utvärdering av kostnaderna för reparation av fiberoptik som gått sönder vid installation, stickprovsmedelvärde $\bar{x}_1 = 65$ och stickprovsvarians $S_1^2 = 25$ för ett stickprov av storlek $n_1 = 21$, och vid användning, stickprovsmedelvärde $\bar{x}_2 = 120$ och stickprovsvarians $S_2^2 = 100$ för ett stickprov av storlek $n_2 = 25$. Målet är styrka påståendet att variansen av kostnaderna för fel vid användning är störst av de två.

a) Formulera passande noll- och alternativhypotes.

b) Använd den givna datan för att test H_0 vid en signifikansnivå på $\alpha = 0.1$.

Lösning:

a) Vi ansätter

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_2^2 > \sigma_1^2$$

b) Sökt: Avgör huruvida vi kan förkasta H_0 eller inte vid $\alpha = 0.1$.

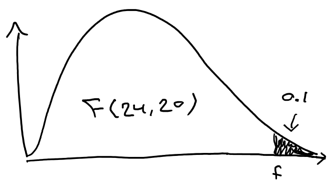
Data: $\bar{x}_1 = 65$, $s_1^2 = 25$, $n_1 = 21$
$\bar{x}_2 = 120$, $s_2^2 = 100$, $n_2 = 25$

$s_2^2 > s_1^2$ så vi väljer att betrakta $\frac{S_2^2}{S_1^2}$ (istället för $\frac{S_1^2}{S_2^2}$). Antag H_0 är sann.

Sats 10.2.1: $\frac{S_2^2}{S_1^2} \sim F(n_2 - 1, n_1 - 1) = F(24, 20)$ $\boxed{\text{Sats 10.2.1 kräver att } \sigma_1^2 = \sigma_2^2}$

$\alpha = 0.1$. Vi vill hitta f sådant att ($F_{24,20} \sim F(24,20)$)

$$P(F_{24,20} \geq f) = 0.1$$



Tabell IX sida 706 ($r_1 = 24$, $r_2 = 20$) ger oss att $f = 1.767$ uppfyller

$$P(F_{24,20} \geq 1.767) = 0.1$$

Så vi förkastar H_0 om $\frac{S_2^2}{S_1^2} > 1.767$. Vi får

$$\frac{S_2^2}{S_1^2} = \frac{100}{25} = 4 > 1.767$$

$\therefore$ Vi förkastar H_0 vid $\alpha = 0.1$

Uppgift 10.14

Följande data har registrerats för en ny och för en gammal laserskannare (enhet: antal skannade streckkoder per sekund)

Ny	Gammal
$n_1 = 61$	$n_2 = 61$
$\bar{X}_1 = 40$	$\bar{X}_2 = 29$
$S_1^2 = 24.9$	$S_2^2 = 22.7$

a) Testa

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 \neq \sigma_2^2$$

med signifikansnivå $\alpha = 0.2$.

b) Bestäm S_p^2.

c) Bestäm ett 90%-konfidensintervall för $\mu_1 - \mu_2$.

d) Verkar det som att den nya skannern läser fler streckkoder i genomsnitt?

e) Antalet streckkoder är diskret och vi har inte rättfärdigat kravet på normalfördelning. Vilken sats rättfärdigar detta?

Lösning:

a) Sökt: Bör vi förkasta H_0 vid $\alpha = 0.2$?

Antag att H_0 är sann. $S_1^2 > S_2^2$, så vi väljer att betrakta $\frac{S_1^2}{S_2^2}$.

H_0 antas vara sann $\rightsquigarrow$ vi kan använda Sats 10.2.1.

Sats 10.2.1: $\frac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1) = F(60,60)$.

Låt $F_{60,60} \sim F(60,60)$.

Vi söker f sådant att

$$P(F_{60,60} \geq f) = \frac{\alpha}{2} = 0.1$$

Sats 10.2.1 kräver normalfördelning

Tabell IX sida 708 ($\gamma_1 = 60, \gamma_2 = 60$) ger oss att $f = 1.395$ uppfyller

$$P(F_{60,60} \geq 1.395) = 0.1$$

Så vi förkastar H_0 om $\frac{S_1^2}{S_2^2} > 1.395$. Vi får att

$$\frac{S_1^2}{S_2^2} = \frac{24.9}{22.7} \approx 1.097 < 1.395$$

$\therefore$ Vi kan inte förkasta H_0 vid $\alpha = 0.2$

b) Sökt: S_p^2.

Def. 10.3.1 ger oss

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \overset{\substack{S_1^2 = 24.9,\ S_2^2 = 22.4 \\ n_1 = 61,\ n_2 = 61}}{=} \frac{60 \cdot 24.9 + 60 \cdot 22.4}{61+61-2} = \boxed{23.65}$$

c) Sökt: 90%-KI för $\mu_1 - \mu_2$.

Från a) får vi att vi inte kan dra slutsatsen att $\sigma_1^2 \neq \sigma_2^2$ (vid $\alpha = 0.2$), så vi antar att $\sigma_1^2 = \sigma_2^2$.

Sats 10.3.1: Ett $100(1-\alpha)$%-KI för $\mu_1 - \mu_2$ ges av

$$P(L_1 \leq \mu_1 - \mu_2 \leq L_2) = 1 - \alpha$$

Sats 10.3.1 kräver att $\sigma_1^2 = \sigma_2^2$

där

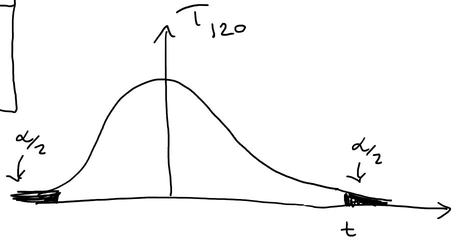
$$P(L_1 \le \mu_1 - \mu_2 \le L_2) = 1 - \alpha$$

där

$$L_1 = \bar{X}_1 - \bar{X}_2 - t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$L_2 = \bar{X}_1 - \bar{X}_2 + t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Sats 10.3.1 kräver att $\sigma_1^2 = \sigma_2^2$
Sats 10.3.1 kräver normalfördelning



och $t_{\alpha/2}$ hör till $T_{n_1+n_2-2}$.

Låt $T \sim T_{n_1+n_2-2} = T_{120}$. Vi söker t sådant att

$$P(T \le t) = 1 - \frac{\alpha}{2} = 0.95.$$

$\gamma = 120$ frihetsgrader finns ej i Tabell VI, så vi använder rader med $\gamma =$ "∞".

Tabell VI sida 700 ($\gamma = \infty$) ger att $t = 1.645$ uppfyller

$$P(T \le 1.645) = 0.95$$

dvs, $t_{\alpha/2} = t = 1.645$. Vi beräknar

$$\bar{X}_1 - \bar{X}_2 = 40 - 29 = 11$$

$$t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \approx 1.45$$

$t_{\alpha/2} = 1.645$
$S_p^2 = 23.65$
$n_1 = 61$
$n_2 = 61$

$$\Rightarrow L_1 \approx 9.55$$

$$L_2 \approx 12.45$$

$\therefore$ Vårt 90%-KI för $\mu_1 - \mu_2$ ges av

$$KI = [9.55, 12.45]$$

d) Ja, eftersom $0 \notin [9.55, 12.45]$ så drar vi slutsatsen att den nya skannern läser fler streckkoder i genomsnitt.

e) Denna argumentation är OK tack vare CLT ($n_1 = 61, n_2 = 61$).

Uppgift 10.24

Följande data har registrerats för en ny och en gammal metod för att förvätska kol (i lämplig enhet)

ny : 1

16.4	12.8	15.4	17.0
17.7	12.2	18.7	
15.9	14.7	19.1	
11.3	14.1	16.5	

gammal : 2

11.1	10.5	10.9	10.1
12.8	13.2	12.6	
12.1	14.5	15.6	
14.2	15.3	14.2	

a) Formulera passande noll- och alternativhypotes för att undersöka huruvida den nya metoden är bättre i genomsnitt vid $\alpha = 0.01$.

b) Utvärdera hypotesen i a) med $\alpha = 0.01$. Skulle du rekommendera den nya metoden istället för den gamla?

Lösning:

a) Vi ansätter (μ_1 är medelvärde för metod 1 och μ_2 är medelvärde för metod 2)

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$ $\quad (\alpha = 0.01)$

b) Sökt: Avgör huruvida vi kan förkasta H_0 vid $\alpha = 0.01$ eller inte.

Först: lika varianser eller ej? Antag normalfördelade data och antag H_0 är sann.
Liknande hypoteser som i 10.14 för att testa detta

$\tilde{H}_0 : \sigma_1^2 = \sigma_2^2$

$\tilde{H}_1 : \sigma_1^2 \neq \sigma_2^2$

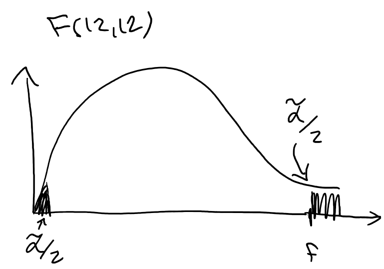
Kan vi förkasta $\hat{H}_0$? Signifikansnivå $\tilde{\alpha}$? Vi väljer $\tilde{\alpha}$ större än α för att minimera vissa problem (se sida 341), vi väljer $\tilde{\alpha} = 0.2$.
Antag $\hat{H}_0$ är sann, så att vi kan använda Sats 10.2.1.
Vi beräknar s_1^2 och s_2^2 (använd Sats 6.3.1)

$s_1^2 \approx 5.89$

$s_2^2 \approx 3.39$

dvs $s_1^2 > s_2^2$ och vi betraktar alltså $\frac{S_1^2}{S_2^2}$.

Sats 10.2.1: $\frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1) = F(12, 12)$. $\boxed{\text{Sats 10.2.1 kräver normalfördelning}}$



Vi söker f sådant att ($F_{12,12} \sim F(12,12)$)

$P(F_{12,12} \geq f) = \frac{\tilde{\alpha}}{2} = 0.1$

Tabell IX sida 705 ($\gamma_1 = 12, \gamma_2 = 12$) ger att $f = 2.147$ uppfyller

$P(F_{12,12} \geq 2.147) = 0.1$

Vi beräknar

$\frac{s_1^2}{s_2^2} \approx \frac{5.89}{3.39} \approx 1.737 < 2.147$,

dvs vi kan <u>inte</u> förkasta $\tilde{H}_0$ vid $\tilde{\alpha} = 0.2$. Eftersom vi <u>inte</u> kan dra slutsatsen att $\sigma_1^2 \neq \sigma_2^2$ så är det <u>OK</u> att använda pooling. (Nu betraktar vi H_0 vs. H_1 igen)
$\Rightarrow$ Vi vill använda Sats 10.3.1, vi behöver S_p^2 och t_α från T - fördelning...

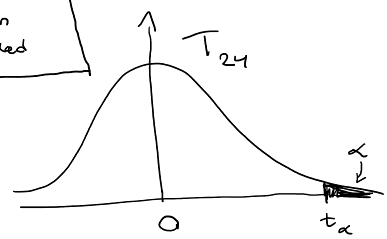
att $\sigma_1 \neq \sigma_2$ så är det OK att använda pooling. (Nu betraktar vi H_0 vs. H_1 igen)

$\Rightarrow$ Vi vill använda Sats 10.3.1, vi behöver S_p^2 och t_α från $T_{n_1+n_2-2}$-fördelning.

Def. 10.3.1 ger oss

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \overset{\substack{s_1^2 \approx 5.89,\ n_1=13 \\ s_2^2 \approx 3.39,\ n_2=13}}{\approx} \frac{12 \cdot 5.89 + 12 \cdot 3.39}{13+13-2} = 4.64$$

Sats 10.3.1 kräver att $\tilde{H}_0$ är sann och normalfördelad data



Vi söker t_α sådant att ($T \sim T_{n_1+n_2-2} = T_{24}$)

$$P(T \leq t_\alpha) = 1-\alpha = 1-0.01 = 0.99$$

Tabell VI sida 699 ($\gamma = 24$) ger oss att $t_\alpha = 2.492$ uppfyller

$$P(T \leq 2.492) = 0.99$$

Vi inför teststatistikan

$$\tilde{T} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{\mu_1=\mu_2}{=} \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Innan experiment: $\tilde{T} \sim T_{n_1+n_2-2} = T_{24}$. Så från ovan ($t_\alpha = 2.492$) har vi att

$$P(\tilde{T} \leq 2.492) = 0.99$$

Data ger oss

$\bar{x}_1 = 15.52$

$\bar{x}_2 = 12.85$

Vi beräknar utfallet av teststatistikan

$$\tilde{T} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{\substack{\bar{x}_1=15.52,\ n_1=13,\ S_p^2 \approx 4.64 \\ \bar{x}_2=12.85,\ n_2=13}}{=} \frac{15.52 - 12.85}{\sqrt{4.64\left(\frac{1}{13} + \frac{1}{13}\right)}} \approx 3.16 > 2.492,$$

dvs vi förkastar H_0 och accepterar H_1 istället. Ja, vi skulle rekommendera den nya metoden.

Uppgift 10.28

En tillverkare har två val av leverantörer för att köpa in komponenter som behöver tåla tryck. Följande data samlas in

leverantör I	leverantör II
$n_1 = 10$	$n_2 = 10$
$\bar{X}_1 = 1350$	$\bar{X}_2 = 1338$
$S_1^2 = 100$	$S_2^2 = 29$

a) Vi söker ett 95% - konfidensintervall för $\mu_1 - \mu_2$. Kan vi poola?

b) Bestäm ett 95% - konfidensintervall för $\mu_1 - \mu_2$.

c) Finns det belägg för att komponenterna från leverantör I i genomsnitt tål högre tryck än komponenterna från leverantör II?

Lösning:

a) Sökt: Kan vi poola?

Vi gör som i uppgift 10.24 och väljer $\tilde{\alpha} = 0.2$. Från uppgift 10.24 får vi att om $\frac{S_1^2}{S_2^2} > f$ så drar vi slutsen att $\sigma_1^2 \neq \sigma_2^2$ (dvs vi kan inte poola), där f uppfyller ($F_{9,9} \sim F(n_1-1, n_2-1) = F(9,9)$)

$$P(F_{9,9} \geq f) = \frac{\tilde{\alpha}}{2} = 0.1$$

Vi behöver anta att data är normalfördelad för denna argumentation

Tabell IX sida 704 ($\gamma_1 = 9, \gamma_2 = 9$) ger oss att $f = 2.440$ uppfyller

$$P(F_{9,9} \geq 2.440) = 0.1$$

Vi beräknar

$$\frac{S_1^2}{S_2^2} = \frac{100}{29} \approx 3.45 > 2.440,$$

dvs vi drar slutsatsen att $\sigma_1^2 \neq \sigma_2^2$ och vi kan alltså inte poola.

b) Sökt: 95%-KI för $\mu_1 - \mu_2$.

Från a): Vi kan inte poola. Sida 347 säger följande

$$T := \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \overset{\text{app.}}{\sim} T_{\lfloor \gamma \rfloor}$$

för något gamma. Vi får en skattning av gamma från Smith-Satterthwaite-formeln enligt följande (se sida 348)

$$\gamma \approx \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$$

Vi beräknar γ

$$\gamma = \frac{(100/10 + 29/10)^2}{\frac{(100/10)^2}{9} + \frac{(29/10)^2}{9}} = \frac{\frac{(100+29)^2}{100}}{\frac{(100/10)^2 + (29/10)^2}{9}} = \frac{129^2 \cdot 9}{100 \cdot [(100/10)^2 + (29/10)^2]} \approx \frac{149769}{10841}$$

$$\approx 13.8$$

Vi avrundar ner till $\lfloor \gamma \rfloor = 13$. Vi söker $t_{\alpha/2}$ ($\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$) sådant att

$$P(T \leq t_{\alpha/2}) = 1 - \frac{\alpha}{2} = 0.975$$

där $T \sim T_{13}$. Tabell VI sida 699 ($\nu = 13$) ger att $t_{\alpha/2} = 2.160$ uppfyller

$$P(T \leq 2.160) = 0.975$$

T-fördelningen är symmetrisk så detta ger

$$P(-2.160 \leq T \leq 2.160) = 0.95$$

Vi sätter in T

$$P\left(-2.160 \leq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \leq 2.160\right) = 0.95$$

Multiplicera allt med $\sqrt{S_1^2/n_1 + S_2^2/n_2}$

$$\Rightarrow P\left(-2.160 \cdot \sqrt{S_1^2/n_1 + S_2^2/n_2} \leq (\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2) \leq 2.160 \cdot \sqrt{S_1^2/n_1 + S_2^2/n_2}\right) = 0.95$$

Subtrahera $\bar{X}_1 - \bar{X}_2$ från alla sidor

$$\Rightarrow P\left(-(\bar{X}_1 - \bar{X}_2) - 2.160\sqrt{S_1^2/n_1 + S_2^2/n_2} \leq -(\mu_1 - \mu_2) \leq -(\bar{X}_1 - \bar{X}_2) + 2.160\sqrt{S_1^2/n_1 + S_2^2/n_2}\right) = 0.95$$

Multiplicera allt med -1

$$\Rightarrow P\left((\bar{X}_1 - \bar{X}_2) + 2.160\sqrt{S_1^2/n_1 + S_2^2/n_2} \geq \mu_1 - \mu_2 \geq (\bar{X}_1 - \bar{X}_2) - 2.160\sqrt{S_1^2/n_1 + S_2^2/n_2}\right) = 0.95$$

$$\Leftrightarrow P(L_1 \leq \mu_1 - \mu_2 \leq L_2) = 0.95$$

där

$$L_1 = (\bar{X}_1 - \bar{X}_2) - 2.160\sqrt{S_1^2/n_1 + S_2^2/n_2}$$

$$L_2 = (\bar{X}_1 - \bar{X}_2) + 2.160\sqrt{S_1^2/n_1 + S_2^2/n_2}$$

Vi beräknar

$$\bar{X}_1 - \bar{X}_2 = 1350 - 1338 = 12$$

$S_1^2 = 100$, $n_1 = 10$
$S_2^2 = 29$, $n_2 = 10$

$$\sqrt{S_1^2/n_1 + S_2^2/n_2} \approx 3.59$$

och

$$L_1 \approx 12 - 2.160 \cdot 3.59 = 4.2456$$

$$L_2 \approx 12 + 2.160 \cdot 3.59 = 19.7544$$

$\therefore$ Vårt 95%-KI för $\mu_1 - \mu_2$ ges approximativt av

$$KI = [4.2, 19.8]$$

c) Ja, eftersom $0 \notin [4.2, 19.8]$

Uppgift 10.43

Ett företag har två leverantörer att välja på. Man tror att leverantör X tar mer betalt än vad leverantör Y tar betalt för liknande produkter. Följande data har registrerats

Produkt	Pris X, $	Pris Y, $
1	$6000 = X_1$	$5900 = Y_1$
2	$575 = X_2$	$580 = Y_2$
3	$15000 = X_3$	$15000 = Y_3$
4	$150\,000 = X_4$	$145000 = Y_4$
5	$76000 = X_5$	$75000 = Y_5$
6	$5650 = X_6$	$5600 = Y_6$
7	$10000 = X_7$	$9975 = Y_7$
8	$850 = X_8$	$870 = Y_8$
9	$900 = X_9$	$890 = Y_9$
10	$3000 = X_{10}$	$2900 = Y_{10}$

Stödjer dessa data misstanken med $\alpha = 0.05$?

Lösning:

Sökt: Avgör om vår data stödjer detta vid $\alpha = 0.05$.

Vi har par av priser (pris hos X och pris hos Y), så det är rimligt att använda "Wilcoxon Signed-Rank test for Paired Observations" (se sida 354). Låt M_X beteckna medianen för X och M_Y beteckna medianen för Y. Vi ansätter

$$H_0: M_X = M_Y$$
$$H_1: M_X > M_Y$$

Antag att H_0 är sann. I detta fall är vi intresserade av

$$|W_-| = \sum_{\substack{\text{alla negativa}\\ \text{ranger}}} |R_i| ,$$

Kom ihåg: $W_+ = \sum_{\substack{\text{alla positiva}\\ \text{ranger}}} R_i$. Om $M_X = M_Y$, då förväntar vi oss att W_- och W_+ är ungefär lika stora. Vi beräknar R_i och $|R_i|$

Produkt	$\lvert X_i - Y_i \rvert$	rang	rang med tecken
1	100	7.5	7.5
2	5	2	-2
3	0	1	-1
4	5000	10	10
5	1000	9	9
6	50	6	6
7	25	5	5
8	20	4	-4
9	10	3	3
10	100	7.5	7.5

Detta ger oss

$$|W_-| \overset{\text{def.}}{=} \sum_{\substack{\text{alla negativa}\\ \text{ranger}}} |R_i| = |-2| + |-1| + |-4| = 7.$$

Hur sannolikt var detta utfall?

Tabell VIII sida 703 (one-sided, $p = \alpha = 0.05$, $n = 10$) ger att

$$P(W_- \geq 11) = 1 - \alpha = 0.95$$

$$\Leftrightarrow P(W_- < 11) = \alpha = 0.05$$

Vi beräknade $W_- = 7 < 11$, vilket är osannolikt under H_0. Vi förkastar alltså H_0 och accepterar istället H_1. Ja, vår data stödjer misstanken om att X tar mer betalt än vad Y gör.