Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Løsningsforslag på eksamen i
SIF5062/SIF5506 Statistikk
Mandag 28. mai 2001

Oppgave 1

a)

$$\begin{aligned} P(Y<12) &= P\left(Z<\frac{12-13}{1{,}5}\right)=P(Z<-0{,}67)=0{,}251 \\ P(11<Y<14) &= P(Y<14)-P(Y<11)=P(Z<0{,}67)-P(Z<-1{,}33) \\ &= 0{,}749-0{,}092=0{,}657. \end{aligned}$$

b) μ=gjennomsnittlig giftkonsentrasjon på havbunnen like ved fabrikken. Vi tester

$$H_0:\mu\le 12 \text{ mot } H_1:\mu>12.$$

Testobservatoren er

$$T=\frac{\overline{Y}-12}{S_Y/\sqrt{n}}$$

som har observert verdi

$$t=\frac{12{,}62-12}{0{,}606/\sqrt{5}}=2{,}29.$$

Den kritiske verdien på nivå 0,05 er $t_{0{,}05,\,4}=2{,}13$. Vi kan dermed forkaste H_0.

c)

$$\begin{aligned} \hat{\beta} &= \frac{\sum(x_i-\bar{x})y_i}{\sum=(x_i-\bar{x})^2}=\frac{\sum x_iy_i-\bar{x}\sum y_i}{\sum(x_i-\bar{x})^2} \\ &= \frac{5245-(550/15)\cdot 160{,}4}{18\,333{,}3}=-0{,}035. \end{aligned}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = \frac{1}{15}[160{,}4 - 20(-0{,}035) \cdot 550] = 11{,}97.$$

Her er $\hat{\alpha}$ estimatet for gjennomsnittlig giftkonsentrasjon på havbunnen like ved fabrikken. Dvs. et estimat for μ i oppgave b.

Estimatet for α er under 12. Da kan vi nødvendigvis ikke konkludere med at $\alpha > 12$. F.eks. får vi en p-verdi på over 0,50.

Dette er en annen konlusjon enn i punkt b. Dette kan skje hvis den lineære modellen ikke er riktig i hele verdiområdet. Et plott av de aktuelle dataene tyder på giftkonsentrasjonen avtar raskere med x når x er liten enn når x er stor.

Oppgave 2

a) Hvis vi tenker oss at alle soldater i en k gruppe blir kontrollert, vil vi ha k forsøk, alle uavhengige, og samme sannsynlighet p for smitte i hvert forsøk. Da er antall smittede binomisk fordelt med parametre k og p. Sannsynligheten for positiv reaksjon i blodprøveblandingen er da

$$1 - P(\text{ingen smittede blant de } k) = 1 - (1-p)^k.$$

b)

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Sett A = "19 Haugen positiv" og B = "k-gruppe positiv". Da er

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{p}{1-(1-p)^k}.$$

c) Den tilfeldige variable Y = "antall grupper som må analyseres på nytt" er binomisk fordelt med forventning $m[1-(1-p)^k]$. Antall analyser X er lik $m + kY$. Dermed får vi

$$\begin{aligned} E(X) &= E(m+kY) = m + kE(Y) \\ &= m + km[(1-(1-p)^k] = mk\left(1 + \frac{1}{k} - (1-p)^k =\right). \end{aligned}$$

Vi må ha at $E(X) < mk$ for at metoden skal lønne seg i det lange løp. Det gir

$$4m\left(1 + \frac{1}{4} - (1-p)^4\right) < 4m,$$

som er oppfylt hvis $p < 1 - \sqrt[4]{1/4} \approx 0{,}29$.

Oppgave 3

a)

$$\begin{aligned} M_X(t) &= E(e^{Xt}) = \int_0^\infty e^{xt}\frac{\theta^4}{6}x^3e^{\theta x}dx = \int_0^\infty \frac{\theta^4}{6}x^3e^{-(\theta-t)x}dx \\ &= \int_0^\infty \frac{\theta^4}{6}x^{\alpha-1}e^{-\frac{x}{\beta}}dx \end{aligned}$$

der $\alpha = 4$ og $\beta = (\theta - t)^{-1}$. I videre utregning benyttes at integranden over er nesten lik den for en Gammafordeling;

$$\begin{aligned} M_X(t) &= \frac{\theta^4}{6}\beta^\alpha\Gamma(\alpha)\int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}dx = \frac{\theta^4}{6}\beta^\alpha\Gamma(\alpha)\cdot 1 \\ &= \frac{\theta^4}{6}(\frac{1}{\theta-t})^4(4-1)! = (\frac{\theta}{\theta-t})^4 = (1-\frac{t}{\theta})^4. \end{aligned}$$

$M_X(t) > 0$ for alle t. Da må vi ha $t < \theta$.

$$\begin{aligned} M_X^{'}(t) &= \frac{4}{\theta}\left(1-\frac{t}{\theta}\right)^{-5}. \\ E(X) &= M_X^{'}(0) = \frac{4}{\theta}. \end{aligned}$$

b)

$$\begin{aligned} L(\theta) &= \prod_{i=1}^n f(x_i). \\ l(\theta) &= \ln L(\theta) = \sum_{i=1}^n \ln f(x_i) \\ &= 4n\ln\theta - n\ln 6 + 3\sum_{i=1}^n \ln x_i - \theta\sum_{i=1}^n x_i. \\ l^{'}(\theta) &= \frac{4n}{\theta} - \sum_{i=1}^n x_i. \\ l^{'}(\hat{\theta}) &= 0 \Rightarrow \hat{\theta} == 20\frac{4n}{\sum_{i=1}^n X_i}. \end{aligned}$$

c)

$$M_{\sum X_i}(t) = \prod_{i=1}^{n} M_X(t) = \left(1 - \frac{t}{\theta}\right)^{-4n}.$$

$$M_{2\theta \sum X_i}(t) = M_{\sum X_i}(2\theta t) = \left(1 - \frac{2\theta t}{\theta}\right)^{-4n} = (1-2t)^{-8n/2}.$$

Vi har (fra tabell) at hvis Y er χ^2-fordelt med $8n$ frihetsgrader, så er

$$M_Y(t) = (1-2t)^{-8n/2} = M_{2\theta \sum X_i}(t).$$

Dermed må $2\theta \sum X_i$ ha samme fordeling som Y.

$$\begin{aligned} P\left(\chi^2_{1-\alpha/2,\, 8n} < 2\theta \sum X_i =< \chi^2_{\alpha/2,\, 8n}\right) &= 1-\alpha \\ P\left(\frac{\chi^2_{1-\alpha/2,\, 8n}}{2\sum X_i} < \theta < \frac{\chi^2_{\alpha/2,\, 8n}}{2\sum X_i}\right) &= 1-\alpha \end{aligned}$$

Et 95% konfidensintervall for θ er da gitt ved

$$\left(\frac{\chi^2_{0{,}975,\, 24}}{2\sum x_i},\ \frac{\chi^2_{0{,}025,\, 24}}{2\sum x_i} =\right) = \left(\frac{12{,}40}{2\cdot 4{,}5},\ \frac{39{,}36}{2\cdot 4{,}5}\right) = (1{,}38,\ \ 4{,}37).$$

Oppgave 4

Det er to ulike måter å tenke på. Den første er at vi trekker fire riktige svar, uten tilbakelegging, fra ni delspørsmål, og registrerer hvor mange av disse som er fra oppgave 1. Da blir X hypergeometrisk fordelt med sannsynlighetsfordeling

$$P(X=x) = \frac{\binom{3}{x}\binom{6}{4-x}}{\binom{9}{4}},\ x = 0,1,2,3.$$

Alternativt kan vi tenke oss at vi trekker tre delspørsmål, uten tilbakelegging, fra ni delspørsmål, hvorav fire er riktige, og registrerer hvor mange riktige vi trekker. Da blir X hypergeometrisk fordelt med sannsynlighetsfordeling

$$P(X=x) = \frac{\binom{4}{x}\binom{5}{3-x}}{\binom{9}{3}},\ x = 0,1,2,3.$$

Det er enkelt å vise at de to sannsynlighetsfordelingene er identiske.