Skilaverkefni 2

Brynjólfur Gauti Jónsson
Þórarinn Jónmundsson

Dæmi 1

Gerum ráð fyrir að meðalverð á gulli hafi verið undanfarið $\mu = 30\$/gr$ og að staðalfrávik verðs á gulli hafi verið $\sigma = 2\$/gr$. Gerum svo ráð fyrir að verð á gulli í næstu viku sé normaldreifð slembistærð með væntigildi $\mu = 30\$/gr$ og staðalfrávik $\sigma = 2\$/gr$.

Dreififall meðalverðs:

$$P(X \leqslant x) = \int_{-\infty}^{x} \frac{1}{\sqrt{8\pi}} e^{-\frac{(X-30)^2}{8}}$$

1. Reiknið líkurnar á því að verð á gulli í næstu viku verði minna en $32\$/gr$.

Lausn: Leysum þetta með R fallinu `pnorm(32, mean = 30, sd = 2)`

Líkurnar eru **84%**

2. Reiknið líkurnar á því að verð á gulli í næstu viku verði meira en $26\$/gr$.

Lausn: Leysum þetta með R fallinu `1 - pnorm(26, mean = 30, sd = 2)`

Líkurnar eru **97.7%**

3. Reiknið líkurnar á því að verð á gulli í næstu viku verði meira en $26\$/gr$ og minna en $32\$/gr$.

Lausn: Leysum þetta með R fallinu `pnorm(32, mean = 30, sd = 2) - pnorm(26, mean = 30, sd = 2)`

Líkurnar eru **82%**

4. Ef við vitum að að verð á gulli í næstu viku verði meira en $30\$/gr$, hverjar eru þá líkurnar á því að það verði meira en $34\$/gr$?

Lausn: Vitum að

$$P(X > A|X > B) = \frac{P(X > A)}{P(X > B)} = \frac{1 - P(X \leqslant A)}{1 - P(X \leqslant B)}$$

Leysum þetta því með R fallinu

`(1 - pnorm(34, mean = 30, sd = 2)) / (1 - pnorm(30, mean = 30, sd = 2))`,

og fáum út að líkurnar séu **5%**

Dæmi 2

Látum $X_i \sim \mathcal{N}(\mu = 3, \sigma^2 = 9), i = 1, ..., 10$ vera óháðar normaldreifðar slembistærðir. Reiknið

$$P(2(X1 + X2 + X3 + X4) < 48)$$

Lausn: Þar sem X_i eru óháðar dreifingar fæst

$$Z = 2\sum_{i=1}^{4} X_i \sim \mathcal{N}(2\sum_{i=1}^{4} \mu_{X_i}, 2\sum_{i=1}^{4} \sigma^2_{X_i}) = \mathcal{N}(24, 72).$$

Auðvelt er að reikna $P(Z < 48)$ með R skipuninni `pnorm(48, 24, 72)` og sjáum að líkurnar eru **63%**

Dæmi 3

Líftími ákveðinna rafhlaðna fylgir normal dreifingu með væntigildi $\mu = 5$ og staðalfrávik $\sigma = 1.5$ vikur. Þegar rafhlaðan deyr er henni samstundis skipt út fyrir nýja sem fylgir sömu dreifingu *(getum hugsað að hún sé frá sama framleiðanda og sé af sömu gerð)*. Metið líkurnar á því að það þurfi að nota 13 rafhlöður eða fleiri á einu ári.

Lausn: Setjum

$$X \sim \mathcal{N}(5, 1.5^2)$$

Gerum ráð fyrir að líftími rafhlaðna sé óháður líftíma annarra rafhlaðna. Ef það skyldi þurfa að nota 13 rafhlöður á einu ári þarf að gilda að $13X \leqslant 52$. Fáum á sama hátt og í *dæmi 2* að

$$Z = 13X \sim \mathcal{N}(45, 13 \cdot 1.5^2)$$

og notum það til að reikna *P(Z < 52)*. R skipunin `pnorm(52, 45, 13 * 1.5^2)` gefur okkur að líkurnar séu 59%

Dæmi 4

Látum $X_1, ..., X_n$ vera einsdreifðar og óháðar slembistærðir þ.a.

$$X_i \sim \mathcal{N}(\mu, \tau), i = i, ..., n$$

þar sem μ er þekkt og τ er óþekkt dreifni (e. variance).

(a) Sýnið að senileikametillinn fyrir τ er

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

.

Lausn: Þéttifall X_i er

$$f(x|\mu,\tau) = \frac{1}{\sqrt{2\pi\tau}}e^{-\frac{(x-\mu)^2}{2\tau}}$$

Sennileikafall fyrir τ er því

$$\mathcal{L}(\tau|x,\mu) = \prod_{i}^{n}\frac{1}{\sqrt{2\pi\tau}}e^{-\frac{(x_i-\mu)^2}{2\tau}}.$$

Einföldum útreikninga með því að taka logrann af sennileikafallinu.

$$\begin{aligned} l(\tau|x,\mu) = \ln(\mathcal{L}(\tau|x,\mu)) &= \sum_{i}^{n}\left[\ln(1) - \ln(\sqrt{2\pi\tau}) - \frac{(x_i-\mu)^2}{2\tau}\right] \\ &= n\cdot 0 - n\cdot\left[\ln(\sqrt{2\pi}) + \ln(\sqrt{\tau})\right] - \frac{1}{2\tau}\sum_{i}^{n}(x_i-\mu)^2 \\ &= -n\cdot\ln(\sqrt{2\pi}) - n\cdot\ln(\sqrt{\tau}) - \frac{1}{2\tau}\sum_{i}^{n}(x_i-\mu)^2 \end{aligned}$$

Þar sem μ er þekkt getum við diffrað lograsennileikafallið með tilliti til τ og fengið samkvæmt fyrstu gráðu skilyrðum

$$\begin{aligned} \frac{dl}{d\tau} &= -n\cdot\frac{1}{\sqrt{\tau}}\frac{1}{2\sqrt{\tau}} + \frac{1}{2\tau^2}\sum_{i}^{n}(x_i-\mu)^2 \\ &= -\frac{n}{2\tau} + \frac{1}{2\tau^2}\sum_{i}^{n}(x_i-\mu)^2 = 0 \\ \rightarrow \frac{n}{2\tau} &= \frac{1}{2\tau^2}\sum_{i}^{n}(x_i-\mu)^2 \\ \rightarrow n &= \frac{1}{\tau}\sum_{i}^{n}(x_i-\mu)^2 \\ \rightarrow \tau n &= \sum_{i}^{n}(x_i-\mu)^2 \\ \rightarrow \tau &= \frac{1}{n}\sum_{i}^{n}(x_i-\mu)^2 \end{aligned}$$

(b) Reiknið bjaga sennileikametilsins $\hat{\tau}$.

Lausn: Þurfum að finna

$$\mathrm{Bias}_\tau[\hat{\tau}] = E[\hat{\tau}] - \tau$$

Reiknun fyrst væntigildi metilsins $\hat{\tau}$. Þar sem meðaltalið μ er þekkt fáum við einfaldlega að

$$E[\hat{\tau}] = E\left[\frac{1}{n}\sum_i^n (x_i - \mu)^2\right] = \tau.$$

Við fáum því að

$$\mathrm{Bias}_\tau[\hat{\tau}] = 0$$

sem kemur heim og saman miðað við forsendur dæmisins. Önnur væri raunin ef þýðismeðaltalið væri ekki þekkt.

Dæmi 5

Úrtak af $n = 56$ athuguðum bómullarsýnum gefur úrtaksmeðaltalið af hlutfallsteygjanleika bómullarinnar sem $\overline{x} = 8.17$ og úrtaksstaðalfrávikið $s = 1.42$. Finnið 95% öryggisbil fyrir sanna meðalhlutfallsteygjanleikann μ.

Lausn: Þar sem við þekkjum ekki staðalfrávik þýðisins finnum við öryggisbilið með *t-dreifingunni*. Stærð úrtaksins er 56 svo fjöldi frígráða er $df = n - 1 = 55$. Vendigildið t_c er því $t_c = 2.00$ fyrir tvíhliða α. Við fáum:

$$\overline{x} \pm t_c \left(\frac{s}{\sqrt{n}} \right) = 8.17 \pm 2 \cdot \frac{1.42}{\sqrt{56}}.$$

95% öryggisbilið fyrir μ er því $(7.79; 8.55)$.

Dæmi 6

Hér höfum við slembiúrtak að stærð $n = 10$ og búið er að reikna fyrir okkur

$$\sum_{i=1}^{x} x_i = 219,0 \text{ og } \sum_{i=1}^{n} x_i^2 = 4949,9$$

.

Hér má gera ráð fyrir að úrtaið komi úr normaldreifðu þýði.

(a) Reiknið punktmat fyrir μ og σ

Lausn: Þar sem úrtakið kemur úr normaldreifðu þýði notum við metla

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{og} \quad s = \sqrt{\frac{1}{n-1}\sum (x_i - \overline{x})^2},$$

fyrir μ og σ. Innsetning gefur að

$$\hat{\mu} = \overline{x} = \frac{1}{10}\cdot 219.0 = 21.9.$$

Reiknum næst

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \overline{x})^2 &= \sum_{i=1}^{n}(x_i^2) - n\overline{x}^2 \\ &= 4949.9 - 10\cdot(21.9)^2 \\ &\approx 153.8.\end{aligned}$$

Þá er $\hat{\sigma} = s$ auðfundinn:

$$s = \sqrt{\frac{1}{9}\cdot 153.8} \approx 4.134.$$

(b) Reiknið 95% öryggisbil fyrir σ

Lausn: Samkvæmt nótunum hans Birgis er $100(1-\alpha)\%$ öryggisbil fyrir σ gefið með

$$\sigma : \left(\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}}\right)$$

Finnum viðeigandi kí-kvaðrat gildi með R skipuninni `qchisq(p = c(0.025, 0.975), df = 9)`:

$$\chi^2_{0.025,9} = 19.023 \quad \text{og} \quad \chi^2_{0.975,9} = 2.700$$

Innsetning gefur þá að 95% öryggisbil σ er

$$\sigma : \left(\sqrt{\frac{9\cdot(4.134)^2}{19.023}}, \sqrt{\frac{9\cdot(4.134)^2}{2.700}}\right) = (2.843, 7.548)$$

Dæmi 7

Líftími $n = 14$ vélhluta var mældur. Summa líftímanna var 1368 klukkustundir. Gerið ráð fyrir að líftímarnir fylgi veldisdreifingu. Þ.e.a.s. ef við látum X_j tákna j-ta líftíma, þá gildir að X_j sé veldisdreifð slembistærð. Í táknmáli ritast það sem

$$X_j \sim \text{Exp}(\lambda)$$

þar sem λ er óþekktur stiki.

(a) Hvert er punktmatið á meðalgildi líftímanna?

Lausn: Þéttifall veldisdreifingar er $f(x) = \lambda e^{-\lambda x}$. Við byrjum á því að finna sennileikafall veldisdreifingarinnar.

$$\mathcal{L}(\lambda) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right).$$

Við tökum lograrnn af $\mathcal{L}(\lambda)$ og fáum

$$l(\lambda) = \ln(\mathcal{L}(\lambda)) = \ln\left(\lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right)\right) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i.$$

Diffrum með tilliti til λ og athugum fyrstu gráðu skilyrði:

$$\frac{d}{d\lambda} ln(L(\lambda)) = \frac{d}{d\lambda}\left(n \ln \lambda - \lambda \sum_{i=1}^{n} x_i\right) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

Þegar jafnan að ofan er leyst fæst að punktmat á λ er

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}.$$

I dæminum er $n = 14$ og summa líftímanna $\sum_{i=1}^{14} x_i = 1368$ svo $\hat{\lambda} = \frac{1368}{14} = \frac{684}{7}$.

(b) Hvert er 95% öryggisbil fyrir meðalgildið?

Lausn: Notfærum okkur að $100(1-\alpha)\%$ öryggsbil fyrir $\hat{\lambda}$ reiknist:

$$\frac{2n}{\hat{\lambda}\chi^2_{(\alpha/2,2n)}} < \frac{1}{\lambda} < \frac{2n}{\hat{\lambda}\chi^2_{(1-\alpha/2,2n)}}$$

Fyrir dæmi verður ójafnan að ofan:

$$\frac{28}{\frac{7}{684} \cdot \chi^2_{(0.025,28)}} < \frac{1}{\lambda} < \frac{28}{\frac{7}{684} \cdot \chi^2_{(0.975,28)}}.$$

Finnum χ^2 gildi með R skipuninni `qchisq(c(0.025, 0.975), df = 28)` og fáum:

$$\frac{28}{\frac{7}{684} \cdot 44.461} < \frac{1}{\lambda} < \frac{28}{\frac{7}{684} \cdot 15.308}.$$

Svo að 95% öryggisbil $\hat{\lambda}$ er $[\mathbf{61.54}, \mathbf{178.73}]$

(c) Reiknið 95% öryggisbil fyrir $P(X > 150)$.

Lausn:

$$P(X > 150) = 1 - P(X \leqslant 150) = 1 - (1 - e^{-150 \cdot \lambda})$$

Setjum inn endapunkta öryggisbilsins úr lið b og fáum

$$e^{-150 \cdot 61.537} = 0.0874 \quad \text{og} \quad e^{-150 \cdot 178.730} = 0.432.$$

Það er, öryggsibilið er: $[\mathbf{0.087}, \mathbf{0.43}]$.

Aukadæmi

Höfum tvo metla d_1 og d_2 til að meta stikann θ. Um metlana eru gefnar eftirfarandi upplýsingar

$$E[d_1] = \theta, Var(d_1) = 6, E[d_2] = \theta + 2, Var(d_2) = 2$$

Hvorn ættum við að velja til að meta stikann θ?

Lausn:

Setjum $\hat{\mu_1} = E[d_1], \hat{\sigma_1}^2 = Var[d_1] = 6, \hat{\mu_2} = E[d_2], \hat{\sigma_2}^2 = Var[d_2] = 2$.

Án þess að vita hvers eðlis dreifingin á bak við θ sé getum við nýtt okkur lögmál stórra talna til að smíða öryggisbil fyrir væntigildin $\hat{\mu}_1$ og $\hat{\mu}_2$ til að nálga $\hat{\theta}$.

Ef við notum $\hat{\mu}_1$ og $\hat{\sigma}_1$ til að smíða öryggisbil fyrir θ fæst eitthvað í námunda við eftirfarandi.

$$\left[\hat{\mu}_1 - x\frac{\sqrt{6}}{\sqrt{n}} \leqslant \hat{\theta} \leqslant \hat{\mu}_1 + x\frac{\sqrt{6}}{\sqrt{n}}\right]$$

$$\hat{\theta} \in \hat{\mu}_1 \pm x\frac{\sqrt{6}}{\sqrt{n}}$$

Hins vegar ef við notum $\hat{\mu}_2$ og $\hat{\sigma}_2$ og þá staðreynd að væntigildið sé línulegur virki getum við nýtt okkur að $\hat{\mu}_2 = \theta + 2$ og fengið:

$$\left[\hat{\mu}_2 - x\frac{\sqrt{2}}{\sqrt{n}} \leqslant \hat{\theta} + 2 \leqslant \hat{\mu}_2 + x\frac{\sqrt{2}}{\sqrt{n}}\right]$$

$$\left[(\hat{\mu}_2 - 2) - x\frac{\sqrt{2}}{\sqrt{n}} \leqslant \hat{\theta} \leqslant (\hat{\mu}_2 - 2) + x\frac{\sqrt{2}}{\sqrt{n}}\right]$$

$$\hat{\theta} \in (\hat{\mu}_2 - 2) \pm \frac{\sqrt{2}}{\sqrt{n}}$$

Öryggisbilið sem smíðað var með μ_2 er þrengra og því áreiðanlegra. Ef hins vegar úrtaksstærð, n, er mjög há myndi þessi munur fara hverfandi.