Matematisk statistik
Matematikcentrum
Lunds universitet

Lösning: 2022–08–26 kl 8^{00}–13^{00}
FMSF30 & FMSF32
Matematisk statistik

---

1. Vi inför beteckningarna A, B, C för öl som kommer från de tre olika bryggerierna, samt M för mörkt öl. Enligt uppgiften har vi följande sannolikheter (lika stor produktion i alla bryggerierna):

$$
\begin{array}{ccc}
P(A) = 0.5 & P(B) = 0.25 & P(C) = 0.25 \\
P(M|A) = 0.3 & P(M|B) = 0.5 & P(M|C) = 1 - 0.25 = 0.75
\end{array}
$$

   (a) Satsen om total sannolikhet ger

$$
\begin{aligned}
P(M) &= P(M|A)P(A) + P(M|B)P(B) + P(M|C)P(C) = \\
&= 0.3 \cdot 0.5 + 0.5 \cdot 0.25 + 0.75 \cdot 0.25 = 0.4625
\end{aligned}
$$

   (b) Bayes formel ger

$$
P(A|M) = \frac{P(A \cap M)}{P(M)} = \frac{P(M|A)P(A)}{P(M)} = \frac{0.3 \cdot 0.5}{0.4625} \approx 0.3243
$$

2. (a) Vi känner igen standard fallet för en binomialfördelning (fixt antal, $n = 10$, oberoende försök med viss sannolikhet, $p = 0.55$ att lyckas). Låt $\xi \in Bin\,(10, 0.55)$ vara antalet frågor som studenten klarar. Sannolikheten att klara provet på första försöket är

$$
P(\xi \geq 6) = 1 - P(\xi < 6) = 1 - P(\xi \leq 5) = 1 - \sum_{k=0}^{5} \binom{10}{k} \cdot 0.55^k \cdot 0.45^{10-k} = 0.5044
$$

   (b) Återigen får vi en binomial fördelning, nu med $n = 10$ studenter, som var och en har $p = 0.5044$ sannolikhet att klara färdighetsprovet på första försöket. Eller, låt $\eta \in Bin\,(10, 0.5044)$ vara antalet studenter som klarar provet på första försöket. Vilket nu ger

$$
P(\eta = 10) = \binom{10}{10} \cdot 0.5044^{10} \cdot (1 - 0.5044)^0 = 0.5044^{10} = 0.001066
$$

   (c) Eftersom varje student gör provet tills de blir godkända, med sannolikhet $p = 0.5044$ att lyckas (bli godkända) i varje försök så känner vi igen en **ffg-fördelning**; vi ska ju räkna antal prov **inklusive** det godkända provet. Låt $Z \in \mathit{ffg}\,(0.5044)$ vara antalet gånger en student gör provet för att bli godkända. Förväntat antal försök är nu:

$$
E(Z) = \frac{1}{p} = \frac{1}{0.5044} = 1.983
$$

3. En stokastisk variabel ξ har följande täthetsfunktion:

$$
f(x) = K \cdot x^2, \quad 0 \leq x \leq 1
$$

   där K är en okänd konstant.

(a) Eftersom tätheten ska integrerar till 1 har vi att

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 K \cdot x^2 dx = \left[\frac{K}{3} \cdot x^3\right]_0^1 = \left(\frac{K}{3} \cdot 1^3 - \frac{K}{3} \cdot 0^3\right) = \frac{K}{3} = 1$$

Vilket ger att $K = 3$.

(b) Sannolikheten blir nu

$$P(\xi > 0.5) = \int_{0.5}^{\infty} f(x)dx = \int_{0.5}^1 K \cdot x^2 dx = \left[\frac{K}{3} \cdot x^3\right]_{0.5}^1 = \left(\frac{K}{3} \cdot 1^3 - \frac{K}{3} \cdot 0.5^3\right) =$$
$$= \left(\frac{8 \cdot K}{3 \cdot 8} - \frac{K}{3 \cdot 8}\right) = \frac{7 \cdot K}{8 \cdot 3} = [K = 3] = \frac{7}{8}$$

(c) Väntevärdet ges av

$$E(\xi) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x \cdot K \cdot x^2 dx = \int_0^1 K \cdot x^3 dx = \left[\frac{K}{4} \cdot x^4\right]_0^1 =$$
$$= \left(\frac{K}{4} \cdot 1^4 - \frac{K}{4} \cdot 0^4\right) = \frac{K}{4} = [K = 3] = \frac{3}{4}$$

4. Vi har två oberoende och normalfördelade stokastiska variabler:
$\xi \in N(3, 0.5)$ och $\eta \in N(-2, 1)$.

(a) Sannolikheten beräknas lämpligen genom att standardisera normal variabeln

$$P(1 < \xi < 3) = P\left(\frac{1-3}{0.5} < \frac{\xi - 3}{0.5} < \frac{3-3}{0.5}\right) = \Phi\left(\frac{3-3}{0.5}\right) - \Phi\left(\frac{1-3}{0.5}\right) =$$
$$= \Phi(0) - \Phi(-4) = \Phi(0) - (1 - \Phi(4)) = 0.5 - (1 - 0.999968) = 0.49997$$

(b) För väntevärde och varians har vi

$$E(2 \cdot \xi - 3 \cdot \eta + 5) = 2 \cdot E(\xi) - 3 \cdot E(\eta) + 5 = 2 \cdot 3 - 3 \cdot (-2) + 5 = 17$$
$$V(2 \cdot \xi - 3 \cdot \eta + 5) = 2^2 \cdot V(\xi) + (-3)^2 \cdot V(\eta) + 0 = 4 \cdot 0.5^2 + 9 \cdot 1 = 10$$

(c) Eftersom linjär kombinationer av normal variabler också är normalfördelade har vi att

$$2 \cdot \xi - 3 \cdot \eta + 5 \in N(17, \sqrt{10})$$

med värdena från b).

5. En lämplig modell är att mängden glass är normalfördelad med samma varians men olika väntevärden.

(a) Skattningar av väntevärde och varians för de två åren blir nu

Förra:

$n_1 = 9$

$\mu_1^* = (59 + 68 + 64 + 70 + 66 + 77 + 72 + 69 + 62)/9 \approx 67.44$

$s_1^2 = ((59 - \mu_1^*)^2 + (68 - \mu_1^*)^2 + \ldots (62 - \mu_1^*)^2)/(9 - 1) \approx 29.53$

Årets:

$n_2 = 7$

$\mu_2^* = (53 + 56 + 55 + 51 + 62 + 58 + 58)/7 \approx 56.14$

$s_2^2 = ((53 - \mu_2^*)^2 + (56 - \mu_2^*)^2 + \ldots (58 - \mu_2^*)^2)/(7 - 1) \approx 13.14$

Eftersom det är en samma varians gör vi en poolad varians skattning

$$s_p^2 = \frac{s_1^2(n_1-1)+s_2^2(n_2-2)}{n_1+n_2-2} = 22.51$$
$$s_p = \sqrt{s_p^2} = 4.74$$

Vi är intresserade av skillnaden mellan de två väntevärdena, osäkerheten i skattningen är

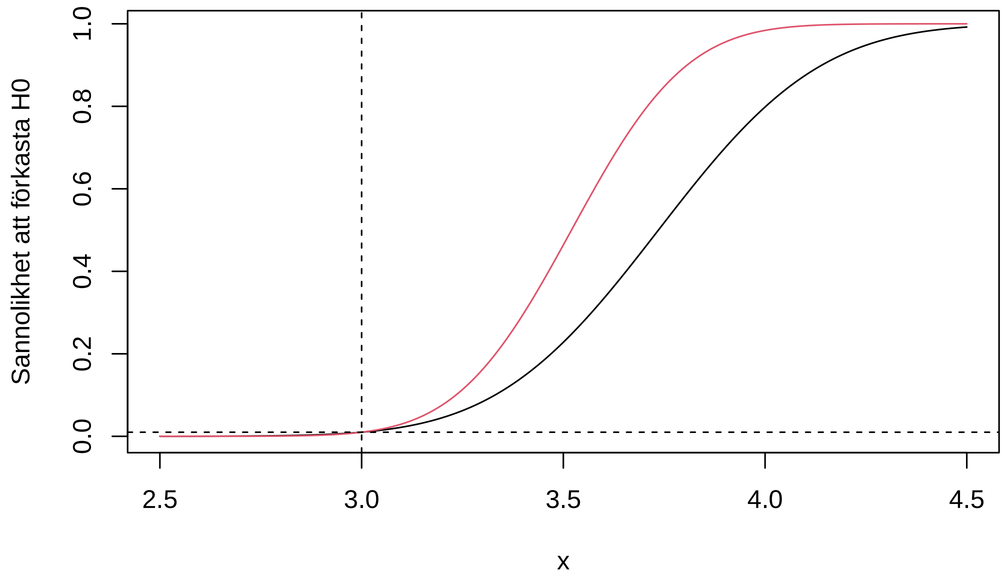
$$V(\mu_2^* - \mu_1^*) = V(\mu_2^*) + V(\mu_1^*) = \tfrac{\sigma^2}{n_2} + \tfrac{\sigma^2}{n_1} = \sigma^2\left(\tfrac{1}{n_2}+\tfrac{1}{n_1}\right)$$

och intervallet blir

$$\begin{aligned} I_{\mu_2-\mu_1} &= \mu_2^* - \mu_1^* \pm t_{0.025}(n_1+n_2-2)s_p\sqrt{\tfrac{1}{n_2}+\tfrac{1}{n_1}} = \\ &= 56.14 - 67.44 \pm 2.1448 \cdot 4.74\sqrt{\tfrac{1}{7}+\tfrac{1}{9}} = \\ &= [-16.43,\ -6.17] \end{aligned}$$

(b) Eftersom intervallet inte innehåller 0 är det en signifikant (på signifikansnivån $\alpha = 5\%$) minskning av glassmängden.

6. (a) Styrkefunktionen måste 1) Öka när värdena närmar sig mot hypotesen, dvs när $\mu > 3$ och 2) Signifikansnivån är definerad som att felaktigt förkasta H0 om H0 sann, alltså måste funktionen anta värdet 0.01 då $\mu = 3$ (den streckade linjen).



(b) När vi ökar antalet mätningar ökar möjligheten att upptäcka en skillnad och kurvan ändras från den svarta (n=10) till den röda (n=20). Notera att värdet i $\mu = 3$ fortfarand måste vara 0.01.