UPPSALA UNIVERSITET
Matematiska institutionen
Fredrik Jonsson

PROV I MATEMATISK STATISTIK
Sannolikhet och Statistik 5 hp
31:e maj 2010

Kortfattade lösningar

1. $P(+ \mid G) = 0.9$, $P(- \mid G^c) = 0.98$ och $P(G) = 0.07$. Detta ger

$$P(+) = P(G)P(+ \mid G) + P(G^c)P(+ \mid G^c) = 0.07 \cdot 0.9 + 0.93 \cdot 0.02 = 0.0816.$$

Med Bayes formel får vi de eftersökta

$$\begin{aligned} P(G \mid +) &= P(+ \mid G) \cdot P(G)/P(+) \approx 77,2\% \\ P(G^c \mid -) &= P(- \mid G^c) \cdot P(G^c)/P(-) \approx 99,2\% \end{aligned}$$

2. (a) Om deltagaren inte byter dörr så är vinst samma sak som att han/hon valt rätt från början. Sannolikheten för detta är 1/4.
   (b) Om deltagaren byter dörr i steg två så är vinst samma sak som att han/hon valt fel i steg 1. Sannolikheten för detta är 3/4.

3. $\mu = \sigma = 1$, $\mu - 2\sigma = -1$, $\mu + 2\sigma = 3$.
   (a) $P(X < -1) = 0$, ty $X \geq 0$.
   (b) $P(X > 3) = e^{-3} \approx 5\%$.

4. Bestäm x så att $P(\bar{X} < x) = 0.9$. Pga. CGS så gäller att $\bar{X}$ är approximativt normalfördelad, $E(\bar{X}) = 70$ och $D(\bar{X}) = 5/\sqrt{20}$. Här efterfrågas en 10%–kvantil för $\bar{X}$. Formler för normalfördelningar ger: $x = 70 + \lambda_{0.1} \cdot 5/\sqrt{20} \approx 71.4$.

5. (a) $I_p = p^* \pm 1.96\sqrt{p^*q^*/n}$. Vi stoppar in $p^* = 0.2$, $q^* = 0.8$ och $n = 50$ vilket ger $\approx [0.09, 0.31]$.
   (b) Samma som (a) men $n = 100$ ger $\approx [0.12, 0.28]$.

6. (a) $p_1^* = 0.47$, $p_2^* = 0.45$ med $X_1 \sim Bin(2000, p_1)$ och $X_2 \sim Bin(2000, p_2)$ ger $I_{p_1 - p_2} = p_1^* - p_2^* \pm 1.96\sqrt{p_1^*q_1^*/2000 + p_2^*q_2^*/2000} \approx [-0.01, 0.05]$. En minskning är inte säkerställd, eftersom värdet 0 (lika stora populationsandelar) innefattas av intervallet.
   (b) $p_1^* = 0.07$, $p_2^* = 0.10$ med $X_1 \sim Bin(2000, p_1)$ och $X_2 \sim Bin(2000, p_2)$ ger på samma sätt $I_{p_2 - p_1} = p_2^* - p_1^* \pm 1.96\sqrt{p_1^*q_1^*/2000 + p_2^*q_2^*/2000} \approx [0.01, 0.05]$. Ökningen är genom detta säkerställd, eftersom intervallet enbart innefattar positiva värden.

7. (a) Betrakta respektive rad som stickprov från respektive population (okända väntevärden och varianser). Konfidensintervall, med kvantilen $t_{0.025}(11)$, blir då: $I_{\mu_1} = [73.5, 80.0]$ respektive $I_{\mu_2} = [66.2, 73.3]$.
   (b) Rimligt är att beräkna en 95% undre konfidensgräns för $\mu_1 - \mu_2$:

$$\bar{x} - \bar{y} - \lambda_{0.05}\sqrt{s_x^2/12 + s_y^2/12} \approx 4.9.$$

   Värdet är markant positivt, så hypotesen kan bekräftas.

8. (a) x =tid och y =genomsnittlig ålder. Ett regressionssamband $y = \alpha + \beta x$ skattas med $\alpha^* \approx 23.8$ och $\beta^* \approx 0.72$. Förklaringsgrad $R^2 \approx 0.977$.
   (b) $\beta^* \approx 0.72$ innebär att genomsnittsåldern stiger med 0.72 år per femårsperiod, dvs. ungefär 1.7 månader per år. Medelåldern år 2025 uppskattas rimligen med $x = 11$ i formeln $y = \alpha^* + \beta^* x$, dvs. 31.7 år.