STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik

MT1011
20 augusti 2020

Lösningar till tentamen i
Sannolikhetslära och statistik för lärare
20 augusti 2020

Examinator: Jan-Olov Persson

Uppgift 1

a) Ett spridningsdiagram (se boken) med bränsleförbrukning på y-axeln och luftryck på x-axeln ska användas.

b) För anpassning av regressionslinjen $y = \hat{a} + \hat{b}x$ med minsta kvadratmetoden beräknas $\hat{a}$ och $\hat{b}$ enligt formelbladet. Regressionslinjen blir $y = 0.952 - 0.09x$ och ritas in i diagrammet (intercept 0.952, lutning -0.09).

c) Uppskattad bränsleförbrukning för lufftryck 2.2 bar blir $0.952-0.09 \cdot 2.2 = 0.754$ liter/mil .

Uppgift 2

a) Att ingen av de fem personerna är född på en vardag betyder att alla fem är födda på en lördag eller söndag och sannolikheten för det är $(2/7)^5 = 0.0019$

b) Antalet av de fem som är födda på en vardag kan beskrivas med en binomialfördelad slumpvariabel $X \sim Bin(n = 5, p = 5/7)$ och sannolikheten för att precis tre är födda på en vardag blir

$$P(X = 3) = \binom{5}{3} \cdot (5/7)^3 \cdot (2/7)^2 = 0.297$$

c) Antaganden: i) Samma sannolikhet, 1/7, för födelse på var och en av veckans dagar (diskret likformig fördelning). ii) Oberoende mellan födelsedagar.

Om man dessutom väljer de fem utan återläggning (vilket är rimligt med tanke på frågans formulering) så måste man också anta att befolkningen är så pass stor att approximation av hypergeometrisk fördelning med binomial kan göras.

Uppgift 3

a) Det ska gälla att $\int_0^c f(x)dx = 1$. Vi får ekvationen

$$1 = \int_0^c \frac{x^3}{4}dx = \left[\frac{x^4}{16}\right]_0^c = \frac{c^4}{16}$$

med lösning $c = 2$.
(Ekvationen har även lösningen $c = -2$ men den är inte aktuell pga villkoret $0 \leq x \leq c$).

b) Med $c = 2$ och $c/2 = 1$ får vi

$$P(X < 1) = \int_0^1 \frac{x^3}{4}dx = \left[\frac{x^4}{16}\right]_0^1 = \frac{1}{16} = 0.0625$$

c) Det som söks är väntevärdet av X, som blir

$$E(X) = \int_0^2 x\frac{x^3}{4}dx = \left[\frac{x^5}{20}\right]_0^2 = \frac{2^5}{20} = 1.6$$

d) Medianen av X, $\tilde{x}$, bestäms med ekvationen

$$0.5 = P(X < \tilde{x}) = \int_0^{\tilde{x}} \frac{x^3}{4}dx = \left[\frac{x^4}{16}\right]_0^{\tilde{x}} = \frac{\tilde{x}^4}{16}$$

som har lösning

$$\tilde{x} = 8^{1/4} \approx 1.68$$

Uppgift 4

a) Låt X vara vikten på ett vitt ägg och Y vikten på ett brunt ägg. Det gäller att $X \sim N(\mu = 55, \sigma = 4)$ och $Y \sim N(\mu = 50, \sigma = 3)$.

$$P(53 \leq X \leq 63) = \phi\left(\frac{63 - 55}{4}\right) - \phi\left(\frac{53 - 55}{4}\right) = 0.669$$

$$P(53 \leq Y \leq 63) = \phi\left(\frac{63-50}{3}\right) - \phi\left(\frac{63-50}{3}\right) = 0.159$$

Cirka 67% av de vita äggen och 16% av de bruna äggen kan förväntas vara medelstora.

Not. $\phi\left(\frac{63-50}{3}\right)$ finns inte med i tabellen och kan sättas till 1.

b) Sannolikheten att det bruna ägget väger mer än det vita om ett av varje väljs kan skrivas som $P(Y > X) \Leftrightarrow P(X - Y < 0)$. Det gäller att $X - Y$ också är normalfördelad men med väntevärde $E(X - Y) = E(X) - E(Y) = 55 - 50 = 5$ och varians $V(X - Y) = V(X) + V(Y) = 4^2 + 3^2 = 25$ om X och Y är oberoende (vilket vi får anta). Sålunda är $X - Y \sim N(\mu = 5, \sigma = 5)$ och

$$P(X - Y < 0) = \phi\left(\frac{0-5}{5}\right) = 1 - \phi(1) = 0.159$$

Uppgift 5

a) Antalet celler i 5 ml, X, antas vara Poissonfördelat med väntevärde $\lambda = 5$ och

$$P(X = 5) = e^{-5}\frac{5^5}{5!} = 0.175$$

b) Antalet celler i 1 ml, X, är Poissonfördelat med väntevärde $\lambda = 1$ och

$$P(X \geq 1) = 1 - P(X = 0) = e^{-1}\frac{1^0}{0!} = 0.632$$

c) Antalet celler i 60 ml, X, är Poissonfördelat med väntevärde $\lambda = 60$ och approximativt gäller att $X \sim N(\mu = 60, \sigma = \sqrt{60})$.

$$P(X > 50) = 1 - P(X \leq 50) = 1 - \phi\left(\frac{50-60}{\sqrt{60}}\right) = 0.90$$

Uppgift 6

Låt p_A och p_B vara sannolikheten att ett frö gror av sort A respektive B. Vi vill testa hypotesen
$H_0 : p_A = p_B \Leftrightarrow p_A - p_B = 0$ mot

$H_1 : p_A \neq p_B \Leftrightarrow p_A - p_B \neq 0$
Om vi väljer signifikansnivå 5% kan det göras med ett 95% konfidensintervall för differensen $p_A - p_B$ som beräknas med formeln (se formelblad)

$$\hat{p}_A - \hat{p}_B \pm 1.96\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$

Med $n_A = n_B = 100$, $\hat{p}_A = 82/100$ och $\hat{p}_B = 75/100$ blir intervallet 0.07 ± 0.113 och eftersom noll ligger i intervallet kan nollhypotesen inte förkastas. Skillnaden i grobarhet mellan de två frösorterna är inte statistiskt signifikant, på nivån $\alpha = 5\%$.