DEL A: ENDAST SVAR SOM ANGES PÅ SVARSBLAD

1. (a) Sätt $p = P(\text{ett visst läkemedel kan expedieras}) = 0.96$. Eftersom läkemedlen är oberoende av varandra blir $P(\text{alla 10 kan expedieras}) = p^{10} = 0.96^{10} = \mathbf{0.665}$.

   (b) Sätt $p_a = P(\text{alla 10 kan expedieras en viss gång}) = 0.665$. Eftersom gångerna är oberoende av varandra blir $P(\text{minst 1 saknas varje gång}) = (1 - p_a)^4 = (1 - 0.665)^4 = \mathbf{0.013}$.

   (c) Sätt $\xi_i =$ ”tiden tills ett restnoterat läkemedel är tillgängligt igen” $\in Exp(\lambda)$ med $\lambda = 1/2$. För ett läkemedel gäller att
   $P(\xi_i \leq 3) = F(3) = \int_0^3 f(x)\,dx = \int_0^3 \frac{1}{2}\,\mathrm{e}^{-x/2}\,dx = 1 - \mathrm{e}^{-3/2} = 0.777$ så att
   $P(\max(\xi_1, \xi_2) > 3) = 1 - P(\max(\xi_1, \xi_2) \leq 3) = 1 - P(\xi_1 \leq 3 \cap \xi_2 \leq 3) = [\text{ober.}] =$
   $= 1 - P(\xi_1 \leq 3) \cdot P(\xi_2 \leq 3) = 1 - 0.777^2 = \mathbf{0.396}$.

   (d) Sätt $\xi =$ ”lufttrycket en junidag” $\in N(1015, 7)$.
   Då blir $P(\xi > 1029) = 1 - \Phi(\frac{1029 - 1015}{7}) = 1 - \Phi(2) = \mathbf{0.023}$.

   (e) Enligt satsen om total sannolikhet gäller att
   $P(\text{regn}) = P(\text{regn} \mid \text{L}) \cdot P(\text{L}) + P(\text{regn} \mid \text{M}) \cdot P(\text{M}) + P(\text{regn} \mid \text{H}) \cdot P(\text{H}) =$
   $= 0.90 \cdot 0.20 + 0.47 \cdot 0.72 + 0.06 \cdot 0.08 = \mathbf{0.523}$.

   (f) Enligt definitionen av betingad sannolikhet gäller att
   $P(\text{kväll} \mid \text{morgon}) = \frac{P(\text{kväll} \cap \text{morgon})}{P(\text{morgon})} = \frac{0.09}{0.19} = \mathbf{0.474}$.

   (g) Sätt $\xi_i =$ ”mängden nederbörd under månad i” för $i = 1, \ldots, n$ där $n = 12$, $E(\xi_i) = \mu = 47.3$ och $D(\xi_i) = \sigma = 29.0$. Då blir $\eta = \sum_{i=1}^{12} \xi_i =$ ”totala mängden nederbörd under ett år”.
   Eftersom $n = 12$ kan anses vara stort och de olika årsnederbördena har samma fördelning och är oberoende[1] av varandra ger CGS att
   $\eta \underset{\sim}{\in} N(\mu n,\, \sigma\sqrt{n}) = N(47.3 \cdot 12,\, 29.0\sqrt{12}) = N(567.6,\, 100.5)$ och
   $P(400 < \eta < 700) = \Phi(\frac{700 - 567.6}{1005}) - \Phi(\frac{400 - 567.6}{100.5}) = \Phi(1.32) - \Phi(-1.67) =$
   $= 0.9066 - (1 - 0.9525) = \mathbf{0.859}$.

   (h) $P(\xi > 100) = 1 - F(100) = 1 - (1 - \mathrm{e}^{-(100/a)^c}) = \mathrm{e}^{-(100/53.0)^{1.7}} = \mathbf{0.053}$.

   (i) Vi söker $x_{0.01}$ där $P(\xi > x_{0.01}) = 1 - F(x_{0.01}) = 1 - (1 - \mathrm{e}^{-(x_{0.01}/a)^c}) = \mathrm{e}^{-(x_{0.01}/a)^c} = 0.01$ vilket ger $x_{0.01} = a \cdot (-\ln 0.01)^{1/c} = 53.0 \cdot (-\ln 0.01)^{1/1.7} = \mathbf{130.1\ mm}$.

   (j) Sätt $\xi =$ ”antal månader med mer än 150 mm nederbörd” $\in Po(\mu)$ och testa H_0: $\mu = 1.2$ mot H_1: $\mu > 1.2$ på signifikansnivån $\alpha = 0.05$. **Ja**, ökningen är signifikant eftersom P-värdet $= P(\xi \geq 5 \,\|\, \xi \in Po(1.2)) = 1 - P(\xi \leq 4) = 1 - 0.992 = \mathbf{0.008} < \alpha = 0.05$.

[1] *Anm.*: I uppgift 2a anges att standardavvikelsen för totala mängden nederbörd under ett år är 106 mm, inte 100.5 som vi får här. Det beror på att månaderna inte är helt oberoende.

DEL B: FULLSTÄNDIGA LÖSNINGAR

2. (a) Sätt ξ_i = ”totala mängden nederbörd år i” för $i = 1, \ldots, n$ där $n = 30$, $E(\xi_i) = \mu$ och $\sigma = 106$. Eftersom $n = 30$ är stort och de olika årsnederbördena har samma fördelning och är oberoende av varandra ger CGS att $\mu^* = \bar{\xi} \underset{\sim}{\in} N(\mu, \frac{\sigma}{\sqrt{n}}) = N(\mu, \frac{106}{\sqrt{30}})$.

   (b) Vi vill bestämma k så att $P(\bar{\xi} > k \,\|\, \mu = 568) = 1 - \Phi(\frac{k - 568}{106/\sqrt{30}}) = 0.05 \Rightarrow$
   $\frac{k - 568}{106/\sqrt{30}} = \lambda_{0.05} = 1.64 \Rightarrow k = 568 + \lambda_{0.05} \cdot \frac{106}{\sqrt{30}} = 600\,\text{mm}.$

   (c) Styrkan är $P(\bar{\xi} > 600 \,\|\, \mu = 655) = 1 - \Phi(\frac{600 - 655}{106/\sqrt{30}}) = 1 - \Phi(-2.86) = \Phi(2.86) =$
   $= 0.9979.$

   (d) Styrkan är $P(\bar{\xi} > 600 \,\|\, \mu = \mu) = 1 - \Phi(\frac{600 - \mu}{106/\sqrt{30}}) > 0.90 \Rightarrow$
   $\frac{600 - \mu}{106/\sqrt{30}} < -\lambda_{0.1} = -1.28 \Rightarrow \mu > 600 + 1.28 \cdot \frac{106}{\sqrt{30}} = 624.5\,\text{mm}.$

3. (a) Eftersom det finns en mycket tydlig systematisk årstidsvariation i medeltemparatur kan man *inte* använda modellen för två oberoende stickprov. Den modellen förutsätter att alla månader har samma medeltemperatur och att de olika månadernas uppmätta värden varierar slumpmässigt och normalfördelat kring totalmedlet. Som man kan se i figuren är det uppenbart inte fallet! Men, eftersom det är samma årstidsvariation på 1700-talet och nu kan man få bort den genom att använda stickprov i par. Förändringen månad i modelleras då som $\xi_i \in N(\Delta, \sigma)$ där Δ är den systematiska förändringen. Skattningarna blir
   $\Delta^* = \bar{x} = 1.40\,^\circ\text{C}$ och $\sigma^* = s = 0.79\,^\circ\text{C}$. Eftersom $\bar{\xi} \in N(\Delta, \frac{\sigma}{\sqrt{12}})$ med medelfelet $d(\Delta^*) = \frac{s}{\sqrt{12}} = \frac{0.79}{\sqrt{12}} = 0.23$ ges konfidensintervallet för den systematiska förändringen Δ av $I_\Delta = (\Delta^* \pm t_{0.025}(12 - 1) \cdot d(\Delta^*)) = (1.40 \pm 2.20 \cdot 0.23) = (0.90,\ 1.90)\,^\circ\text{C}$.

   (b) Eftersom H_0: $\Delta = 0$ inte ligger i konfidensintervallet kan H_0 förkastas till förmån för H_1: $\Delta \neq 0$. Ja, det finns en systematisk temperaturförändring.

4. (a) $\beta^* = \frac{S_{xy}}{S_{xx}} = \frac{41.77}{68.92} = 0.61$ där $\beta^* \in N(\beta, \frac{\sigma}{\sqrt{S_{xx}}})$ med
   $\sigma^* = s = \sqrt{\frac{Q_0}{n-2}} = \sqrt{\frac{S_{yy} - S_{xy}^2/S_{xx}}{n-2}} = \sqrt{\frac{13.0}{24-2}} = 0.77$ och
   medelfelet $d(\beta^*) = \frac{s}{\sqrt{S_{xx}}} = \frac{0.77}{\sqrt{68.92}} = 0.09$ så att
   $I_\beta = (\beta^* \pm t_{0.025}(24 - 2) \cdot d(\beta^*)) = (0.61 \pm 2.07 \cdot 0.09) = (0.41,\ 0.80)\,^\circ\text{C/h}.$

   (b) Med $\alpha^* = \bar{y} - \beta^* \cdot \bar{x} = \frac{440.7}{24} - 0.61 \cdot \frac{183.8}{24} = 13.7$ ges ett konfidensintervall för linjen
   $\mu_0 = \alpha + \beta_0 \cdot x_0 = \alpha + \beta \cdot 2$ av $I_{\mu_0} = (\alpha^* + \beta^* \cdot x_0 \pm t_{0.025}(24 - 2) \cdot s\sqrt{\frac{1}{24} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}) =$
   $= (13.7 + 0.61 \cdot 2 \pm 2.07 \cdot 0.77\sqrt{\frac{1}{24} + \frac{(2 - \frac{183.8}{24})^2}{68.92}}) = (14.9 \pm 2.07 \cdot 0.55) =$
   $= (13.8,\ 16.1)\,^\circ\text{C}.$

(c) Eftersom konfidensintervallet för β_3 innehåller 0 är det ingen signifikant skillnad i lutning, dvs 1 extra soltimme ger inte signifikant annorlunda temperaturändring på vintern jämfört med på sommaren. Eftersom konfidensintervallet för β_2 inte innehåller 0 är det däremot en signifikant systematisk skillnad mellan vinter och sommar, så att det på vintern är ungefär 13 grader kallaren än på sommaren, för samma antal soltimmar.

(d) Med $x_1 = 2$, $x_2 = 1$ och $x_3 = x_1 \cdot x_2 = 2$ får vi $\beta_0^* + \beta_1^* \cdot 2 + \beta_2^* \cdot 1 + \beta_3^* \cdot 2 =$
$= 13.7 + 0.6 \cdot 2 - 12.9 \cdot 1 - 0.7 \cdot 2 = 0.6\,^\circ\mathrm{C}$.

5. (a) Sätt ξ = "antal dagar MD8h överstiger $120\,\mu\mathrm{g/m}^3$ på ett år" och $p_0 = P(\xi = 0)$ som uppskattas med $p_0^* = \dfrac{10}{19} = 0.53$.

(b) Om vi betraktar data som en diskret, känd, fördelning får vi

$$\mu = E(\xi) = \sum_{x=0}^{7} x \cdot P(\xi = x) = 0 \cdot \frac{10}{19} + \cdots + 7 \cdot \frac{1}{19} = 1.21.$$

$$\sigma^2 = V(\xi) = E(\xi^2) - \mu^2 = \sum_{x=0}^{7} x^2 \cdot P(\xi = x) - 1.21^2 = 0^2 \cdot \frac{10}{19} + \cdots + 7^2 \cdot \frac{1}{19} - 1.21^2 = 3.43.$$

$$\sigma = \sqrt{3.43} = 1.85.$$

Anm. Om man vill se data som ett stickprov istället för en känd fördelning kommer variansskattningen att behöva justeras lite för att divideras med $n - 1 = 18$ istället för med $n = 19$: $(\sigma^2)^* = \dfrac{19}{18} \cdot 3.43 = 3.62$ och $\sigma^* = \sqrt{3.62} = 1.90$.

(c) Sätt η = "antal år man uppfyller normen" $\in Bin(5,\ p_0) = Bin(5,\ 0.53)$.
Det ger $P(\eta = 2) = \dbinom{5}{2} 0.53^2 (1 - 0.53)^3 = 0.29$.

(d) Vi söker n så att $P(\sum_{i=1}^{n} \xi_i \geq 42) \geq 0.90$ där $E(\xi) = \mu = 1.21$ och $D(\xi) = \sigma$ enligt 5b.

Eftersom n kommer att vara stort kan vi använda CGS och få att $\sum_{i=1}^{n} \xi_i \underset{\sim}{\in} N(\mu n,\ \sigma\sqrt{n})$ och

$P(\sum_{i=1}^{n} \xi_i \geq 42) \approx 1 - \Phi(\dfrac{42 - \mu n}{\sigma\sqrt{n}}) \geq 0.90 \Leftrightarrow \Phi(\dfrac{42 - \mu n}{\sigma\sqrt{n}}) \leq 0.10$, vilket ger villkoret

$$\frac{42 - \mu n}{\sigma\sqrt{n}} \leq -\lambda_{0.1} = -1.28 \Rightarrow$$

$$\sqrt{n} \geq \frac{\lambda_{0.1} \cdot \sigma}{2 \cdot \mu} + \sqrt{\left(\frac{\lambda_{0.1} \cdot \sigma}{2 \cdot \mu}\right)^2 + \frac{42}{\mu}} = \begin{cases} 6.95 & \text{med } \sigma = 1.85, \\ 6.98 & \text{med } \sigma^* = 1.90. \end{cases}$$

Det innebär att man måste mäta i minst $n \geq \begin{cases} 6.95^2 = 48.3 \\ 6.98^2 = 48.7 \end{cases} = 49$ år. Heltal! Och stort.

Slut!