1. (a) Vi söker p så att $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$ Med $\mathsf{P}(A \cup B) = 0.6, \mathsf{P}(A) = 0.4$ och $\mathsf{P}(B) = p$ får vi $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B) \Rightarrow 0.6 = 0.4 + p - 0.4p \Rightarrow 0.6p = 0.2 \Rightarrow p = \frac{1}{3}$

   (b) Vi har $\mathsf{E}(X) = (0 + 1 + 2 + 3)/4 = 3/2$ och
$$\mathsf{V}(X) = \sum_{k=0}^{3}(k - \frac{3}{2})^2 \frac{1}{4} = \frac{(\frac{3}{2})^2 + (\frac{1}{2})^2 + (\frac{1}{2})^2 + (\frac{3}{2})^2}{4} = \frac{5}{4}$$

2. Låt X vara antalet burkar med fel vikt för den nya maskinen och låt Y vara antalet burkar med fel vikt för den gamla maskinen. Vi har X ~Bin$(1000, p_1)$, Y ~Bin$(1000, p_2)$ och vill testa $H_0 : p_1 = p_2$ mot $H_1 : p_1 < p_2$. Vi har $p_1^* - p_2^* = 13/1000 - 21/1000 = -1/125$ och $p^* = (13 + 21)/2000 = 17/1000$. Vi bildar teststorheten
$$T = \frac{p_1^* - p_2^*}{\sqrt{p^*(1 - p^*)(\frac{1}{1000} + \frac{1}{1000})}}$$
som under H_0 är approximativt $\mathsf{N}(0, 1)$-fördelad. Vi har $T_{obs} = -1.3838$ och $z_{0.05} = 1.6449$. Eftersom $|T_{obs}| < z_{0.05}$ kan vi ej förkasta H_0.

3. (a) Vi har $3X - 2Y \sim \mathsf{N}(3 \cdot 3 - 2 \cdot 4, 3^2 \cdot 2^2 + 2^2 \cdot 3^2) = \mathsf{N}(1, 72)$. Alltså är $\mathsf{P}(3X - 2Y < 3) = \Phi(\frac{3-1}{\sqrt{72}}) = \Phi(0.2357) = 0.5932$.

   (b) Om $X \sim \mathsf{N}(250, 10^2)$ är vikten på en konserv så är $p = \mathsf{P}(X > 260) = 1 - \Phi((260 - 250)/10) = 1 - \Phi(1) = 0.1587$. Låt Y vara antalet konserver av totalt fem som väger mer än 260 gram. Vi har då Y ~Bin$(5, p)$ och får den sökta sannolikheten som
$$\begin{aligned}\mathsf{P}(Y \geq 2) &= 1 - \mathsf{P}(Y \leq 1) = 1 - \mathsf{P}(Y = 0) - \mathsf{P}(Y = 1)\\ &= 1 - \binom{5}{0} p^0(1 - p)^5 - \binom{5}{1} p^1(1 - p)^4\\ &= 1 - (1 - p)^5 - 5p(1 - p)^4 = 1 - 0.4216 - 0.3975 = 0.181\end{aligned}$$

4. (a) Vi har
$$\int_0^{100} f(x)dx = \frac{1}{C}\int_0^{100} 100 - xdx = \frac{5000}{C}$$
och ska alltså välja $C = 5000$.

   (b) Vi har
$$\mathsf{E}(X) = \int_0^{100} xf(x)dx = \frac{1}{5000}\int_0^{100} x(100 - x)dx = \frac{1}{5000}\frac{100^3}{6} = \frac{100}{3}$$

5. (a) Vi ansätter regressionsmodellen $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ där ε_i antas vara oberoende $\mathsf{N}(0, \sigma^2)$-fördelade. Skattningen av parametrarna ges av $\beta_1^* = S_{xy}/S_{xx} = 0.43/1.97 = 0.2183$, $\beta_0^* = \bar{y} - \beta_1^*\bar{x} = 0.85 - 0.2183 \cdot 3.0 = 0.1952$. Skattningen av σ^2 fås av $s^2 = Q_0/18 = (S_{yy} - \frac{S_{xy}^2}{S_{xx}})/18 = (0.32 - 0.43^2/1.97)/18 = 0.0126$.

   (b) Vi söker $\mu_Y^*(2.5) = \beta_0^* + 2.5\beta_1^* = 0.7409$. Vi har $t_{\alpha/2}(18) = 2.1$ och får ett tvåsidigt konfidensintervall med $\alpha = 0.05$ som
$$\begin{aligned}I_{\mu_Y^*(2.5)} &= \left(\beta_0^* + 2.5\beta_1^* \pm t_{\alpha/2}(18)s\sqrt{\frac{1}{n} + \frac{(2.5 - \bar{x})^2}{S_{xx}}}\right) = (0.7409 \pm 2.1 \cdot 0.1121 \cdot 0.4206)\\ &= (0.7409 \pm 0.099) = (0.6418, 0.8399)\end{aligned}$$

(c) Om strömstyrkan ökar med 2A så ökar vätskeflödet med $2\beta_1$. En punktskattning av ökningen ges av $2\beta_1^* = 0.4365$. För att göra ett intervall för $2\beta_1$ behöver vi $\mathsf{E}(2\beta_1^*) = 2\beta_1$ och $\mathsf{V}(2\beta_1^*) = 4\mathsf{V}(\beta_1^*) = 4\sigma^2/S_{xx}$. Ett konfidensintervall fås därför på samma sätt som för β_1 men där vi skalar skattning och standardavvikelse med två:

$$I_{2\beta_1} = (2\beta_1^* \pm t_{\alpha/2}(18) 2\frac{s}{\sqrt{S_{xx}}}) = (0.101, 0.772)$$

6. Låt M vara medianen för halten CO_2. Vi vill testa $H_0 : M = 500$ mot $H_1 : M > 500$. Eftersom vi har ett ensidigt test är den intressanta teststorheten $Q_- =$ antalet observationer mindre än 500. Vi har tre mätningar under 500 och får därför p-värdet för teckentestet som

$$\begin{aligned} p &= \mathsf{P}(X \le 3 | X \sim \text{Bin}(13, 0.5)) = \mathsf{P}(X=0) + \mathsf{P}(X=1) + \mathsf{P}(X=2) + \mathsf{P}(X=3) \\ &= \binom{13}{0} 0.5^{13} + \binom{13}{1} 0.5^{13} + \binom{13}{2} 0.5^{13} \binom{13}{3} 0.5^{13} \\ &= 0.5^{13}(1 + 13 + 78 + 286) = \frac{378}{8192} = 0.0461 \end{aligned}$$

Eftersom $p < 0.05$ kan vi förkasta med $\alpha = 0.05$.

7. (a) Antag att vi har observationerna $x_1, \ldots, x_n$. ML-skattaren av μ ges av det värde som maximierar likelihood-funktionen $L(\mu) = \prod_{i=1}^n f(x_i)$ eller log-likelihooden

$$\log L(\mu) = \sum_{i=1}^n \ln f(x_i) = \sum_{i=1}^n \left(-(\log x_i - \mu)^2 - \log x_i - \log \sqrt{\pi} \right)$$

Vi deriverar med avseende på μ och får

$$\frac{\partial \log L(\mu)}{\partial \mu} = \sum_{i=1}^n -2(\log x_i - \mu) = 0 \Rightarrow \sum_{i=1}^n \log x_i = n\mu$$

Så ML-skattaren ges av $\mu^* = \frac{1}{n} \sum_{i=1}^n \log x_i$

(b) Vi har

$$\mathsf{E}(\log X) = \int_0^\infty \frac{\log(x)}{x\sqrt{\pi}} e^{-(\log(x)-\mu)^2} dx$$

Vi gör variabelbytet $y = \log(x)$ och får

$$\mathsf{E}(\log X) = \int_{-\infty}^\infty \frac{y}{\sqrt{\pi}} e^{-(y-\mu)^2} dy = \int_{-\infty}^\infty y f(y) dy$$

där $f(y) = \frac{1}{\sqrt{\pi}} e^{-(y-\mu)^2}$ är täthetsfunktionen till en $\mathsf{N}(\mu, 0.5)$-fördelning. Eftersom $\int_{-\infty}^\infty y f(y) dy$ är definitionen av väntevärdet för en fördelning med täthetsfunktion $f(y)$ så får vi alltså att $\mathsf{E}(\log X) = \mu$.

(c) Från (b) har vi att $\mathsf{E}(\log X) = \mu$. Skattningens väntevärde blir därför

$$\mathsf{E}(\mu^*) = \mathsf{E}\left(\frac{1}{n} \sum_{i=1}^n \log X_i \right) = \frac{1}{n} \sum_{i=1}^n \mathsf{E}(\log X_i) = \frac{1}{n} n\mu = \mu$$

och ML-skattningen av μ är alltså väntevärdesriktig.