1. (a) Vi har $\mathsf{P}(A|B) = P(A \cap B)/\mathsf{P}(B)$. Vidare har vi att

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B),$$

vilket ger $\mathsf{P}(A \cap B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cup B) = 0.4 + 0.3 - 0.5 = 0.2$.
Vi får därför att $\mathsf{P}(A|B) = 0.2/0.3 = 2/3$.

(b) Låt A_i beteckna händelsen att en bil kommer från model i där $i = 1, 2, 3$. Låt K vara händelsen att krockkudden utlöst av misstag. Vill beräkna $P(A_1|K)$. Vi kan antingen använda Bayes sats eller direkt räkna ut sannolikheterna från grunden. Vi tar det senare alternativet. Vi har att

$$\mathsf{P}(A_1|K) = \frac{\mathsf{P}(A_1 \cap K)}{\mathsf{P}(K)}.$$

Vi har även att $\mathsf{P}(A_1 \cap K) = \mathsf{P}(K|A_1)\mathsf{P}(A_1)$ Satsen om total sannolikhet ger dessutom att

$$\mathsf{P}(K) = \sum_{i=1}^{3} P(K|A_i)\mathsf{P}(A_i) = 3 \cdot 10^{-5} \cdot 0.6 + 6 \cdot 10^{-5} \cdot 0.3 + 20 \cdot 10^{-5} \cdot 0.1 = 5.6 \cdot 10^{-6}.$$

Sätter vi samman detta fås

$$\mathsf{P}(A_1|K) = \frac{\mathsf{P}(K|A_1)\mathsf{P}(A_1)}{\sum_{i=1}^{3} P(K|A_i)\mathsf{P}(A_i)} = \frac{3 \cdot 10^{-5} \cdot 0.6}{5.6 \cdot 10^{-5}} = 0.3214.$$

(c) Sätt upp fördelningsfunktionen för Y :

$$\begin{aligned} F_Y(y) &= \mathsf{P}(Y \leq y) = \mathsf{P}(\sqrt{X} \leq y) = \mathsf{P}(X \leq y^2) \\ &= F_X(y^2) = \begin{cases} 0 & y < 0 \\ y^2 & 0 \leq y < 1 \\ 1 & y > 1 \end{cases} \end{aligned}$$

I sista steget har vi använt att $F_X(x) = x$ om $0 \leq x \leq 1$ för en R(0,1)-fördelning. Vi kan nu bestämma tätheten för genom att derivera map y. Vi får då att

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ 2y & 0 \leq y \leq 1 \\ 0 & y > 1 \end{cases}$$

2. Enligt centrala gränsvärdessatsen så är 100 kräftors sammanlagda vikt approximativt normalfördelad. Låt X_i vara vikten i hg för kräfta i där $i = 1, 2, \ldots, 100$. Låt $Y = \sum_{i=1}^{100} X_i$ Vill beräkna $\mathsf{P}(Y < 40)$ Nu är enligt CGS

$$Y \underset{\sim}{\in} N(\mathsf{E}Y, \mathsf{D}Y).$$

Vi har att

$$\begin{aligned} \mathsf{E}[Y] &= \mathsf{E}\left[\sum_{i=1}^{100} X_i\right] \overset{\text{ober.}}{=} \sum_{i=1}^{100} \mathsf{E}[X_i] \overset{\text{lika förd.}}{=} 100\mathsf{E}[X_1], \\ \mathsf{V}[Y] &= \mathsf{E}\left[\sum_{i=1}^{100} X_i\right] \overset{\text{ober.}}{=} \sum_{i=1}^{100} \mathsf{E}[X_i] \overset{\text{lika förd.}}{=} 100\mathsf{V}[X_1]. \end{aligned}$$

Vidare får vi att

$$\begin{aligned}
\mathsf{E}[X_1] &= \int_{0.2}^{2} x\frac{1-x}{0.32} = \frac{7}{15} \\
\mathsf{E}[X_1^2] &= \int_{0.2}^{2} x^2\frac{1-x}{0.32} = \frac{19}{75}
\end{aligned}$$

$$V[X_1] = \mathsf{E}[X_1^2] - \mathsf{E}[X_1]^2 = \frac{19}{75} - \left(\frac{7}{15}\right)^2 = \frac{8}{225}$$

Vi får då att

$$Y \underset{\sim}{\in} N\left(100\frac{7}{15},\ 10\frac{\sqrt{8}}{15}\right).$$

Vilket ger att

$$\mathsf{P}(Y > 40) = \mathsf{P}\left(\frac{Y - 100\frac{7}{15}}{10\frac{\sqrt{8}}{15}} < \frac{40 - 100\frac{7}{15}}{10\frac{\sqrt{8}}{15}}\right) \approx \Phi(-3.5) = 1 - \Phi(3.5) = 2.3 \cdot 10^{-4}$$

3. (a) Vi har två oberoende stickprov $x_1, \ldots, x_{10}$, och $y_1, \ldots, y_{10}$, och vill testa

$$\begin{aligned}
H_0 : \mu_1 &= \mu_2, \\
H_1 : \mu_1 &> \mu_2.
\end{aligned}$$

$\bar{x} = 1.143$ och $\bar{y} = 0.949$. $\bar{x} - \bar{y}$ är en observation av $\bar{X} - \bar{Y} \in N(\mu_1 - \mu_2, \sigma\sqrt{\frac{1}{10} + \frac{1}{10}})$.

Testkvantitet: $u = \frac{\bar{x}-\bar{y}}{\sigma\sqrt{\frac{1}{10}+\frac{1}{10}}} = \frac{0.194}{\sqrt{0.004}} = 3.07$.

Förkasta H_0 om $u > \lambda_{0.01} = 2.33$. H_0 kan alltså förkastas på nivån 0.01.

Alternativt kan vi göra ett ensidigt nedåt begränsat konfidens intervall för $\mu_1 - \mu_2$ enligt

$$I_{\mu_1-\mu_2} = [\bar{x} - \bar{y} - \lambda_{0.01}\sigma\sqrt{\frac{1}{10} + \frac{1}{10}}, \infty] = [0.194 - 2.33 \cdot 0.0632, \infty] = [0.0467, \infty]$$

Intervallet täcker ej över noll och vi kan alltså förkasta H_0 på nivån 0.01.

(b) Vi vill räkna ut styrkan då $\mu_1 - \mu_2 = 0.2$.

$$\begin{aligned}
h(0.2) &= P(H_0 \text{ förkastas om } \mu_1 - \mu_2 = 0.2) = P\left(\frac{\bar{X} - \bar{Y}}{\sigma\sqrt{\frac{1}{10} + \frac{1}{10}}} > \lambda_{0.01} \,|\mu_1 - \mu_2 = 0.2\right) \\
&= P\left(\frac{\bar{X} - \bar{Y} - 0.2}{\sigma\sqrt{\frac{1}{10} + \frac{1}{10}}} > \lambda_{0.01} - \frac{0.2}{\sigma\sqrt{\frac{1}{10} + \frac{1}{10}}}\right) = 1 - \Phi\left(\lambda_{0.01} - \frac{0.2}{\sigma\sqrt{\frac{1}{10} + \frac{1}{10}}}\right) \\
&= 1 - \Phi\left(2.33 - \frac{0.2}{\sqrt{0.02}\sqrt{\frac{1}{10} + \frac{1}{10}}}\right) \\
&\approx 1 - \Phi(-0.83) = \Phi(0.83) \approx 0.7967
\end{aligned}$$

4. (a) Vi beräkna $\mathsf{C}(X, Y) = \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y]$. På grund av symmetri har vi att $\mathsf{E}[X] = \mathsf{E}[Y]$. Vi börjar med att beräkna $\mathsf{E}[X]$:

$$\begin{aligned}
\mathsf{E}[X] &= \int_0^1 \int_0^1 x\left(\frac{5}{4} - xy\right) dydx = \int_0^1 x\left[\left(\frac{5}{4}y - x\frac{y^2}{2}\right)\right]_{y=0}^{y=1} dx \\
&= \int_0^1 x\left(\frac{5}{4} - \frac{x}{2}\right) dx = \left[x^2\frac{5}{8} - \frac{x^3}{6}\right]_0^1 = \frac{5}{8} - \frac{1}{6} = \frac{11}{24}
\end{aligned}$$

Vi fortsätter med $\mathsf{E}[XY]$

$$\begin{aligned}
\mathsf{E}[XY] &= \int_0^1\int_0^1 xy\left(\frac{5}{4}-xy\right)dydx = \int_0^1 x\left[\left(\frac{5}{8}y^2 - x\frac{y^3}{3}\right)\right]_{y=0}^{y=1}dx \\
&= \int_0^1 x\left(\frac{5}{8}-\frac{x}{3}\right)dx = \left[x^2\frac{5}{16}-\frac{x^3}{9}\right]_0^1 = \frac{5}{16}-\frac{1}{9} = \frac{29}{144}.
\end{aligned}$$

Vi får då att

$$\mathsf{C}(X,Y) = \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] = \frac{29}{144} - \frac{121}{576} = \frac{-5}{576}.$$

(b) Enligt formelsamling så är

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

Vi har att

$$f_X(x) = \int f_{X,Y}(x,y)dy = \begin{cases} \int_0^1 \frac{5}{4} - xydy = \frac{5}{4} - \frac{x}{2} & 0 < x < 1 \\ 0 & \text{f.ö.} \end{cases}$$

vilket ger att

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{\frac{5}{4}-xy}{\frac{5}{4}-\frac{x}{2}} & 0<x<1,\ 0<y<1 \\ 0 & \text{f.ö.} \end{cases}$$

(c) Eftersom $\mathsf{C}(X,Y) \neq 0$ så är X och Y beroende. Alternativt kan vi se att $f_{Y|X=x}(y)$ beror på x så X och Y kan därför inte vara oberoende.

5. Vi har följande summor och kvadratsummor

$$\begin{aligned}
n &= 9, \quad \bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{1}{9}\cdot 72 = 8, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{1}{9}\cdot 85.5 = 9.5 \\
S_{xx} &= \sum_{i=1}^n (x_i-\bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n}\left(\sum_{i=1}^n x_i\right)^2 = 636 - \frac{1}{9}\cdot 72^2 = 60 \\
S_{yy} &= \sum_{i=1}^n (y_i-\bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{1}{n}\left(\sum_{i=1}^n y_i\right)^2 = 889.7 - \frac{1}{9}\cdot 85.5^2 = 77.45 \\
S_{xy} &= \sum_{i=1}^n (x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^n x_iy_i - \frac{1}{n}\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right) = \\
&= 751.7 - \frac{1}{9}\cdot 72\cdot 85.5 = 67.7
\end{aligned}$$

(a) Regressionsparametrarna skattas med

$$\begin{aligned}
\beta^* &= \frac{S_{xy}}{S_{xx}} = \frac{67.7}{60} = 1.1283 \\
\alpha^* &= \bar{y} - \beta^*\bar{x} = 9.5 - 1.1283\cdot 8 = 0.4733 \\
(\sigma^2)^* &= s^2 = \frac{1}{n-2}Q_0 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right) = \frac{1}{7}\left(77.45 - \frac{67.7^2}{60}\right) = 0.1517 \\
\sigma^* &= \sqrt{0.1517} = 0.3895
\end{aligned}$$

(b) Vi vill alltså göra ett kalibreringsintervall för x_0 då vi observerat $y = 10.4$ x_0 kan lösas ut ur $y = \alpha^* + \beta^* x_0$ och blir

$$x_0^* = \frac{y - \alpha^*}{\beta^*} = \frac{10.4 - 0.4733}{1.1283} \approx 8.8$$

Från formelsamlingen har vi att

$$\begin{aligned} I_{x_0} &= x_0^* \pm t_{p/2}(9-2)s\sqrt{1 + \frac{1}{n} + \frac{(x_0^* - \bar{x})^2}{S_x x}} = 8.8 \pm 2.36 \cdot 0.3895\sqrt{1 + \frac{1}{9} + \frac{(8.8-8)^2}{60}} \\ &\approx [7.83,\ 9.77] \end{aligned}$$

(c) Om vi betraktar y_i som observationer av $Y_i \in N(\beta x_i, \sigma)$ fås MK-skattningen av β genom att minimera $Q(\beta)$ enligt

$$\begin{aligned} Q(\beta) &= \sum_{i=1}^{n}(y_i - \beta x_i)^2 \\ \frac{dQ}{d\beta} &= -2\sum_{i=1}^{n}(y_i - \beta x_i)x_i = -2\sum_{i=1}^{n} x_i y_i + 2\beta\sum_{i=1}^{n} x_i^2 = 0 \quad \Longrightarrow \\ \beta_{MK}^* &= \frac{\sum_1^n x_i y_i}{\sum_1^n x_i^2} = \frac{751.7}{636} = 1.1819 \end{aligned}$$

6. (a) Vi vill beräkna ML skattningen av ϑ med hjälp både stickproven. Vi sätter upp likelihooden

$$\begin{aligned} L(\vartheta, x, y) &= \prod_{i=1}^{10} p_{X_i}(x_i) \cdot \prod_{k=1}^{30} p_{y_k}(y_k) \\ &= \prod_{i=1}^{10} \frac{\vartheta^{x_i}}{x_i!}e^{-\vartheta} \cdot \prod_{k=1}^{30} \frac{(3\vartheta)^{y_k}}{y_k!}e^{-3\vartheta} \end{aligned}$$

Vi bildar nu log-likelihood funktionen

$$\ln(L(\vartheta, x, y)) = \sum_{i=1}^{10}(x_i \ln(\vartheta) - \ln(x_i!) - \vartheta) + \sum_{k=1}^{30}(x_k \ln(3\vartheta) - \ln(x_k!) - 3\vartheta).$$

Vi deriverar map på ϑ och får

$$\frac{\partial}{\partial\vartheta}\ln(L(\vartheta, x, y)) = \sum_{i=1}^{10}(\frac{x_i}{\vartheta} - 1) + \sum_{k=1}^{30}(\frac{x_k}{\vartheta} - 3) = \frac{1}{\vartheta}(\sum_{i=1}^{10} x_i + \sum_{k=1}^{30} y_k) - 100$$

Vi sätter lika med noll och löser ut ϑ vilket ger

$$\vartheta_{ML}^* = \frac{\sum_{i=1}^{10} x_i + \sum_{k=1}^{30} y_k}{100} = \frac{106}{100} = 1.06$$

Vi vill nu beräkna väntevärde och varians för skattningen

$$\begin{aligned} \mathsf{E}[\vartheta_{ML}^*] &= \mathsf{E}\left[\frac{\sum_{i=1}^{10} X_i + \sum_{k=1}^{30} Y_k}{100}\right] \\ &= \frac{\sum_{i=1}^{10} \mathsf{E}[X_i] + \sum_{k=1}^{30} \mathsf{E}[Y_k]}{100} \\ &\overset{\text{formel saml}}{=} \frac{\sum_{i=1}^{10} \vartheta + \sum_{k=1}^{30} 3\vartheta}{100} = \frac{(10+90)\vartheta}{100} = \vartheta \end{aligned}$$

Vilket ger att skattningen är väntevärdesriktig.
Vi fortsätter med variansen

$$
\begin{aligned}
\mathsf{V}[\vartheta^*_{ML}] &= \mathsf{V}\left[\frac{\sum_{i=1}^{10} X_i + \sum_{k=1}^{30} Y_k}{100}\right] \\
&\overset{\text{ober}}{=} \frac{\sum_{i=1}^{10} \mathsf{V}[X_i] + \sum_{k=1}^{30} \mathsf{V}[Y_k]}{100^2} \\
&\overset{\text{formel saml}}{=} \frac{\sum_{i=1}^{10} \vartheta + \sum_{k=1}^{30} 3\vartheta}{100^2} = \frac{(10+90)\vartheta}{100^2} = \frac{\vartheta}{100} = 0.01\vartheta
\end{aligned}
$$

(b) För att beräkna MK-skattningen ska vi minimera

$$
\begin{aligned}
Q(\vartheta, x, y) &= \sum_{i=1}^{10} (x_i - \mathsf{E}[X_i])^2 + \sum_{k=1}^{30} (y_k - \mathsf{E}[Y_k])^2 \\
&= \sum_{i=1}^{10} (x_i - \vartheta)^2 + \sum_{k=1}^{30} (y_k - 3\vartheta)^2.
\end{aligned}
$$

Vi deriverar map på ϑ och får

$$
\begin{aligned}
\frac{\partial}{\partial \vartheta} Q(\vartheta, x, y) &= \sum_{i=1}^{10} -2(x_i - \vartheta) + \sum_{k=1}^{30} -6(y_k - 3\vartheta) \\
&= -2\left(\sum_{i=1}^{10} x_i + 3\sum_{k=1}^{30} y_k\right) + 2\vartheta(10 + 270).
\end{aligned}
$$

Vi sätter lika med noll och löser ut ϑ vilket ger

$$
\vartheta^*_{MK} = \frac{\sum_{i=1}^{10} x_i + 3\sum_{k=1}^{30} y_k}{280} = \frac{13 + 3 \cdot 93}{280} = 1.0429.
$$

Vi vill nu beräkna väntevärde och varians för skattningen

$$
\begin{aligned}
\mathsf{E}[\vartheta^*_{MK}] &= \mathsf{E}\left[\frac{\sum_{i=1}^{10} X_i + 3\sum_{k=1}^{30} Y_k}{280}\right] \\
&= \frac{\sum_{i=1}^{10} \mathsf{E}[X_i] + 3\sum_{k=1}^{30} \mathsf{E}[Y_k]}{280} \\
&\overset{\text{formel saml}}{=} \frac{\sum_{i=1}^{10} \vartheta + 3\sum_{k=1}^{30} 3\vartheta}{100} = \frac{(10+270)\vartheta}{280} = \vartheta
\end{aligned}
$$

Vilket ger att skattningen är väntevärdesriktig.
Vi fortsätter med variansen

$$
\begin{aligned}
\mathsf{V}[\vartheta^*_{MK}] &= \mathsf{V}\left[\frac{\sum_{i=1}^{10} X_i + 3\sum_{k=1}^{30} Y_k}{280}\right] \\
&\overset{\text{ober}}{=} \frac{\sum_{i=1}^{10} \mathsf{V}[X_i] + \sum_{k=1}^{30} 9\mathsf{V}[Y_k]}{280^2} \\
&\overset{\text{formel saml}}{=} \frac{\sum_{i=1}^{10} \vartheta + \sum_{k=1}^{30} 27\vartheta}{280^2} = \frac{(10+810)\vartheta}{280^2} = \vartheta\frac{41}{3920} \approx 0.0105\vartheta
\end{aligned}
$$

(c) Från (a) och (b) ser vi att både ML och MK skattningarna är väntevärdesriktiga dock har ML-skattningen en aning lägre varians så den är därför att föredra.
