---

1. (a) Låt A_i beteckna händelsen att Maskin i fungerar, för $i = 1, 2, 3$ och låt A beteckna händelsen att precis två maskiner fungerar. A kan inträffa på tre olika sätt, antingen $A_1 \cap A_2 \cap A_3^c$ (dvs att Maskin 1 och Maskin 2 fungerar, men Maskin 3 inte fungerar) eller $A_1 \cap A_2^c \cap A_3$ eller $A_1^c \cap A_2 \cap A_3$. Dessa tre händelser är disjunkta och vi får därför att den sökta sannolikheten är

$$\begin{aligned}
\mathsf{P}(A) &= \mathsf{P}((A_1 \cap A_2 \cap A_3^c) \cup (A_1 \cap A_2^c \cap A_3) \cup (A_1^c \cap A_2 \cap A_3)) \\
&= \mathsf{P}(A_1 \cap A_2 \cap A_3^c) + \mathsf{P}(A_1 \cap A_2^c \cap A_3) + \mathsf{P}(A_1^c \cap A_2 \cap A_3) \\
&= 0.92 \cdot 0.96 \cdot (1 - 0.98) + 0.92 \cdot (1 - 0.96) \cdot 0.98 + (1 - 0.92) \cdot 0.96 \cdot 0.98 \\
&= 0.017664 + 0.036064 + 0.075264 \\
&= 0.128992
\end{aligned}$$

(b) Låt X beteckna den totala kapaciteten per minut, vi har du

$$\begin{aligned}
\mathsf{P}(X = 0) &= \mathsf{P}(A_1^c \cap A_2^c \cap A_3^c) = (1 - 0.92) \cdot (1 - 0.96) \cdot (1 - 0.98) = 0.000064 \\
\mathsf{P}(X = 2) &= \mathsf{P}(A_1^c \cap A_2 \cap A_3^c) = (1 - 0.92) \cdot 0.96 \cdot (1 - 0.98) = 0.001536 \\
\mathsf{P}(X = 3) &= \mathsf{P}(A_1^c \cap A_2^c \cap A_3) = (1 - 0.92) \cdot (1 - 0.96) \cdot 0.98 = 0.003136 \\
\mathsf{P}(X = 5) &= \mathsf{P}(A_1 \cap A_2^c \cap A_3^c) + \mathsf{P}(A_1^c \cap A_2 \cap A_3) \\
&\quad = (1 - 0.92) \cdot 0.96 \cdot 0.98 + 0.92 \cdot (1 - 0.96) \cdot (1 - 0.98) = 0.076 \\
\mathsf{P}(X = 7) &= \mathsf{P}(A_1 \cap A_2 \cap A_3^c) = 0.92 \cdot 0.96 \cdot (1 - 0.98) = 0.017664 \\
\mathsf{P}(X = 8) &= \mathsf{P}(A_1 \cap A_2^c \cap A_3) = 0.92 \cdot (1 - 0.96) \cdot 0.98 = 0.036064 \\
\mathsf{P}(X = 10) &= \mathsf{P}(A_1 \cap A_2 \cap A_3) = 0.92 \cdot 0.96 \cdot 0.98 = 0.865536
\end{aligned}$$

Det sökta väntevärdet ges nu av

$$\begin{aligned}
\mathsf{E}(X) &= 0 \cdot \mathsf{P}(X = 0) + 2 \cdot \mathsf{P}(X = 2) + 3 \cdot \mathsf{P}(X = 3) + 5 \cdot \mathsf{P}(X = 5) + 7 \cdot \mathsf{P}(X = 7) \\
&\quad + 8 \cdot \mathsf{P}(X = 8) + 10 \cdot \mathsf{P}(X = 10) \\
&= 9.46.
\end{aligned}$$

En enklare lösning är att inse att maskinerna är oberoende, så om vi låter X_i beteckna det förväntade antalet förpackningar som maskin i producerar under en minut så har vi att

$$\mathsf{E}(X) = \mathsf{E}(X_1) + \mathsf{E}(X_2) + \mathsf{E}(X_3) = 0.92 \cdot 5 + 0.96 \cdot 2 + 0.98 \cdot 3 = 9.46$$

2. (a) Log-likelihooden ges av

$$L(\mathbf{x}; \theta) = \sum_{i=1}^{n} \left(3 \log x_i - 4 \log \theta - \log(6) - \frac{x_i}{\theta} \right) = 3 \sum_{i=1}^{n} \log x_i - 4n \log \theta - n \log(6) - \sum_{i=1}^{n} \frac{x_i}{\theta}$$

Vi deriverar med avseende på θ:

$$\frac{\partial}{\partial \theta} L(\mathbf{x}; \theta) = -4n \frac{1}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i$$

Slutligen sätter vi $\frac{\partial}{\partial \theta} L(\mathbf{x}; \theta) = 0$, löser ut θ, och får

$$\theta^* = \frac{1}{4n} \sum_{i=1}^{n} x_i$$

(b) Om X_i har täthetsfunktion $f(x)$ har vi

$$\begin{aligned}
\mathsf{E}(X_i) &= \int_0^\infty x\frac{x^3}{6\theta^4}\exp\left(-\frac{x}{\theta}\right)dx = \frac{1}{6\theta^4}\int_0^\infty x^4\exp\left(-\frac{x}{\theta}\right)dx \\
&= [\text{variabelbyte } y = x/\theta] = \frac{\theta}{6}\int_0^\infty y^4\exp\left(-y\right)dy = \frac{\theta}{6}4! = 4\theta
\end{aligned}$$

Alltså har vi

$$\mathsf{E}(\theta^*) = \frac{1}{4n}\sum_{i=1}^n \mathsf{E}(X_i) = \frac{1}{4n}\sum_{i=1}^n 4\theta = \theta,$$

så skattaren är väntevärdesriktig.

3. (a) Låt X_i vara vikten på äpple i och låt $Y = \sum_{i=1}^6 X_i$ vara den totala vikten. Vi har $Y \sim N(6\cdot 130, 6\cdot 20^2)$ och får

$$\mathsf{P}(Y \geq 800) = 1 - \mathsf{P}(Y < 800) = 1 - \Phi(\frac{800-780}{\sqrt{6\cdot 20^2}}) = 1 - \Phi(0.408) = 0.3416.$$

(b) Låt $Y_k = \sum_{i=1}^k X_i$. Vi söker k så att

$$\mathsf{P}(Y_k \geq 800) = 1-\mathsf{P}(Y_k < 800) = 1-\Phi(\frac{800-130\cdot k}{\sqrt{k\cdot 20^2}}) = 0.95 \Rightarrow \frac{800-130\cdot k}{\sqrt{k\cdot 20^2}} = \Phi^{-1}(0.05).$$

Enligt tabell har vi $\Phi^{-1}(0.05) = -1.6449$, och har alltså $\frac{800-130\cdot k}{\sqrt{k\cdot 20^2}} = -1.6449$. k ges därför av lösningen till ekvationen $130k - 1.6449\cdot 20\sqrt{k} - 800 = 0$. Löser vi denna ekvation får vi att $k = 6.80$, och vi behöver alltså ta minst 7 äpplen för att nå den önskade sannolikheten.

4. (a) Skattningarna ges av $\beta_1^* = S_{xy}/S_{xx} = 0.0002159201$, $\beta_0^* = \bar{y} - \beta_1^*\bar{x} = -26.74022$, och $s = \sigma^* = \sqrt{(S_{yy} - S_{xy}^2/S_{xx})/(n-2)} = \sqrt{45.63838/317} = \sqrt{0.1439696} = 0.3794$.

(b) Vi vill testa $H_0 : \beta_1 = 0$ mot $H_1 : \beta_1 \neq 0$ på nivå $\alpha = 0.01$. Vi beräknar ett konfidensintervall för β_1,

$$\begin{aligned}
I_{\beta_1} &= \left[\beta_1^* \pm t_{\alpha/2}(317)\frac{s}{\sqrt{S_{xx}}}\right] = [0.0002159201 \pm 2.5914\cdot 8.35\cdot 10^{-6}] \\
&= [0.0001942,\ 0.0002375]
\end{aligned}$$

Eftersom detta intervall inte täcker noll kan vi förkasta H_0.

(c) Det förväntade värdet ges av $\mu_Y(300000) = \beta_0^* + 300000\beta_1^* = 38.03582$. Prediktionsintervallet ges av

$$I_{\mu_Y(300000)} = \left(38.03582 \pm t_{\alpha/2}(317)s\sqrt{1 + \frac{1}{319} + \frac{(300000-\bar{x})^2}{S_{xx}}}\right) = [36.638,\ 39.434]$$

5. (a) Låt x_F och x_E beteckna mätningarna hos kvinnorna för och efter studien. På motsvarande sätt låt y_F och y_E beteckna mätningarna hos männen. Vi beräknar först stickprovsmedelvärden och stickprovsvarianser hos datan för respektive grupp. Resultaten sammanfattas i följande tabell:

	Kvinnor		Män	
	Väntevärde	Varians	Väntevärde	Varians
Före	133.3333	13.06667	135.2	10.7
Efter	131.6667	9.466667	131.8	7.7

Vi antar att $x_F \sim \mathsf{N}(\mu_{K_F}, \sigma^2_{K_F})$ och $x_F \sim \mathsf{N}(\mu_{M_F}, \sigma^2_{M_F})$ och vill testa

$$H_0 : \sigma_{K_F} = \sigma_{M_F}$$
$$H_1 : \sigma_{K_F} \neq \sigma_{M_F}$$

på nivå $\alpha = 0.05$. Vi beräknar teststorheten $T_{obs} = s^2_{K_F}/s^2_{M_F} = 13.06667/10.7 = 1.221184$, som under H_0 är $F(5,4)$-fördelad. Enligt formelsamligen är $F_{\alpha/2}(5,4) = 9.36$ och eftersom $T_{obs} < F_{\alpha/2}(5,4)$ kan vi inte förkasta H_0.

(b) Vi har samma modell som i (a) men antar att $\sigma_{K_F} = \sigma_{M_F}$ och vill nu testa

$$H_0 : \mu_{K_F} = \mu_{M_F}$$
$$H_1 : \mu_{K_F} \neq \mu_{M_F}$$

på nivå $\alpha = 0.05$. Den poolade variansskattningen ges av

$$s_p^2 = \frac{5s^2_{K_F} + 4s^2_{M_F}}{9} = \frac{43.6333}{9} = 12.01481$$

Vi beräknar teststorheten

$$T_{obs} = \frac{\bar{x}_F - \bar{y}_F}{s_p\sqrt{\frac{1}{6} + \frac{1}{5}}} = -0.8893496$$

Under H_0 är teststorheten $t(9)$-fördelad, och enligt formelsamligen är $t_{\alpha/2}(9) = 2.2622$. Eftersom $|T_{obs}| < t_{\alpha/2}(9)$ kan vi inte förkasta H_0.

(c) Vi beräknar skillnaderna i blodtryck för varje person:

	kvinnor						män				
Person nr	1	2	3	4	5	6	7	8	9	10	11
Skillnad	-2	-3	2	4	-8	-3	-6	-2	-4	0	-5

Enligt modellerna ovan bör skillnaden för kvinnor vara $\mathsf{N}(\Delta_K, \sigma^2)$-fördelad och skillnaden för män vara $\mathsf{N}(\Delta_M, \sigma^2)$-fördelad. Vi vill nu testa

$$H_0 : \Delta_K = \Delta_M$$
$$H_1 : \Delta_K \neq \Delta_M$$

på nivå $\alpha = 0.05$. Vi har $\bar{D}_K = -1.66667$ $\bar{D}_M = -3.4$. En poolad skattning av σ^2 ges av

$$s_p^2 = \frac{5s^2_{D_F} + 4s^2_{D_M}}{9} = \frac{112.5333}{9} = 12.5037$$

Teststorheten är nu

$$T_{obs} = \frac{\bar{D}_K - \bar{D}_M}{s_p\sqrt{\frac{1}{6} + \frac{1}{5}}} = 0.809519$$

Under H_0 är teststorheten $t(9)$-fördelad, och enligt formelsamligen är $t_{\alpha/2}(9) = 2.2622$. Eftersom $|T_{obs}| < t_{\alpha/2}(9)$ kan vi inte förkasta H_0 i detta fall heller.

6. (a) Låt oss kalla vilken maskin som används för Faktor A, och vilken sirap som används för Faktor B. Vi har ett försök med tvåsidig indelning och ansätter därför modellen

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk},$$

där $\varepsilon_{ijk} \sim \mathsf{N}(0, \sigma^2)$, $i = 1, 2, 3$ betecknar vilken maskin som används, $j = 1$ om sirap av Märke A används och $j = 2$ om sirap av Märke B används. Slutligen betecknar $k = 1, 2, 3$ de tre mätningarna per försöksuppställning.

(b) De medelvärden vi behöver är först medelvärdena för alla faktorkombinationer:

$$\bar{y}_{11} = 3.27 \quad \bar{y}_{21} = 3.34 \quad \bar{y}_{31} = 6.223333, \bar{y}_{12} = 5.593333 \quad \bar{y}_{22} = 7.68 \quad \bar{y}_{32} = 9.12.$$

Vi behöver också det totala medelvärdet och det totala medelvärdet för varje faktor:

$$\begin{aligned}
\bar{y}_{\cdot 1} &= (\bar{y}_{11} + \bar{y}_{21} + \bar{y}_{31})/3 = 4.277778, \\
\bar{y}_{\cdot 2} &= (\bar{y}_{12} + \bar{y}_{22} + \bar{y}_{32})/3 = 7.464444, \\
\bar{y}_{1\cdot} &= (\bar{y}_{11} + \bar{y}_{12})/2 = 4.431667, \\
\bar{y}_{2\cdot} &= (\bar{y}_{21} + \bar{y}_{22})/2 = 5.51, \\
\bar{y}_{3\cdot} &= (\bar{y}_{31} + \bar{y}_{32})/2 = 7.671667, \\
\bar{y} &= (\bar{y}_{\cdot 1} + \bar{y}_{\cdot 2})/2 = 5.871111.
\end{aligned}$$

(c) Vi kan använda medelvärdena från (a) för att beräkna följande kvadratsummor

$$\begin{aligned}
SS_A &= 6\left((\bar{y}_{1\cdot} - \bar{y})^2 + (\bar{y}_{2\cdot} - \bar{y})^2 + (\bar{y}_{3\cdot} - \bar{y})^2\right) = 32.66641, \\
SS_B &= 9\left((\bar{y}_{\cdot 1} - \bar{y})^2 + (\bar{y}_{\cdot 2} - \bar{y})^2\right) = 45.6968, \\
SS_E &= \sum_{ijk}(y_{ijk} - \bar{y}_{ij})^2 = \sum_{ij}(\sum_{k} y_{ijk}^2 - 3\bar{y}_{ij}^2) = 9.208533, \\
SS_{Tot} &= \sum_{ijk}(y_{ijk} - \bar{y})^2 = \sum_{ijk} y_{ijk}^2 - 18\bar{y}^2 = 90.81118, \\
SS_{AB} &= SS_{Tot} - SS_E - SS_A - SS_B = 3.239433
\end{aligned}$$

(d) Baserat kvadratsummorna från (b) kan vi nu fylla i alla fält i variansanalystabellen och får

Variation	Kvadratsumma	Frihetsgrader	Medelkvadrat	Teststorhet
Faktor A	32.66641	2	16.33321	21.28444
Faktor B	45.6968	1	45.6968	59.54929
Faktor AB	3.239433	2	1.619717	2.110716
Residual	9.208533	12	0.7673778	
Total	90.81118	17		

(e) Vi vill testa $H_0 : AB = 0$ mot $H_1 : AB \neq 0$. Med $\alpha = 0.05$ får vi från formelsamlingen att det kritiska värdet för teststorheten är $F_{0.05}(2, 12) = 3.88$. Eftersom $F = 2.110716 < 3.88$ kan vi inte förkasta H_0, och vi har alltså ingen samspelseffekt.

(f) Vi vill testa $H_0 : A = 0$ mot $H_1 : A \neq 0$ samt Vi vill testa $H_0 : B = 0$ mot $H_1 : B \neq 0$.. För Faktor A är det kritiska värdet för teststorhetern $F_{0.05}(2, 12) = 3.88$ och eftersom $F = 21.28444 > 3.88$ kan vi förkasta H_0. För Faktor B har vi istället det krtiska värdet $F_{0.05}(1, 12) = 4.47$ men även i detta fall har vi $F = 59.54929 > 4.47$ och kan förkasta H_0.