1. (a) Vi beräknar först $\mathsf{P}(B) = \mathsf{P}(B|A)\mathsf{P}(A) + \mathsf{P}(B|A^c)\mathsf{P}(A^c) = 0.2 \cdot 0.02 + 0.026 \cdot 0.98 = 0.0295$. Den sökta sannolikheten ges av Bayes formel,

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(B|A)\mathsf{P}(A)}{\mathsf{P}(B)} = \frac{0.2 \cdot 0.02}{0.0295} = 0.1356$$

(b) Nej, eftersom $\mathsf{P}(B|A) \neq \mathsf{P}(B)$.

(c) Låt A, B och C beteckna händelserna att respektive komponent går sönder. Vi söker

$$\mathsf{P}(A \cap B \cap C) = \mathsf{P}(B \cap C|A)\mathsf{P}(A) = \mathsf{P}(C|B, A)\mathsf{P}(B|A)\mathsf{P}(A) = 0.5 \cdot 0.2 \cdot 0.02 = 0.002.$$

2. (a) Täthetsfunktionen för X fås genom att integrera $f(x, y)$ över y:

$$f(x) = \int_0^1 f(x, y)dy = \frac{12}{19}\left[y + x^2y + \frac{xy^2}{2}\right]_0^1 = \frac{12}{19}(1 + x^2 + \frac{x}{2})$$

för $0 \leq x \leq 1$.

(b) Den sökta sannoliheten fås som

$$\begin{aligned}\mathsf{P}(X > Y) &= \int_0^1 \int_0^x f(x, y)dydx = \frac{12}{19}\int_0^1 \left[y + x^2y + \frac{xy^2}{2}\right]_0^x dx \\ &= \frac{12}{19}\int_0^1 x + x^3 + \frac{x^3}{2}dx = \frac{12}{19}\left[\frac{x^2}{2} + \frac{3x^4}{8}\right]_0^1 \\ &= \frac{3 \cdot 7}{19 \cdot 2} \approx 0.5526.\end{aligned}$$

3. (a) Väntevärdet ges av

$$\sum_{i=13}^{18} if(i) = 13 \cdot 0.01 + 14 \cdot 0.04 + 15 \cdot 0.05 + 16 \cdot 0.1 + 17 \cdot 0.3 + 18 \cdot 0.5 = 17.14.$$

Vidare har vi

$$\sum_{i=13}^{18} i^2 f(i) = 13^2 \cdot 0.01 + 14^2 \cdot 0.04 + 15^2 \cdot 0.05 + 16^2 \cdot 0.1 + 17^2 \cdot 0.3 + 18^2 \cdot 0.5 = 295.08.$$

Med hjälp av detta beräknas variansen som

$$\sum_{i=13}^{18} i^2 f(i) - \left(\sum_{i=13}^{18} if(i)\right)^2 = 295.08 - 17.14^2 = 1.3004.$$

(b) Låt X_i vara antalet personer i vagn nummer i. Antalet personer som åker under en timme är $Y = \sum_{i=1}^{60} X_i$. Enligt centrala gränsvärdessatsen gäller att Y är approximativt $\mathsf{N}(60 \cdot 17.41, 60 \cdot 1.3004)$. Vi söker

$$\begin{aligned}\mathsf{P}(Y > 1035) &= 1 - \mathsf{P}(Y \leq 1035) \approx 1 - \Phi\left(\frac{1035 - 1028.4}{\sqrt{60 \cdot 1.3004}}\right) \\ &= 1 - \Phi(0.75) = 1 - 0.7734 = 0.2266.\end{aligned}$$

4. (a) Låt p vara sannolikheten att en student lyckas. Det total antalet lyckade försök är då $\mathrm{Bin}(62, p)$.

   (b) Momentmetoden bygger på att beräkna så många teoretiska moment hos fördelningen som vi har okända parametrar. Sedan beräknas motsvarande stickprovsmoment från datan. Parametrarna skattas genom att lösa ekvationssystemet som fås genom att sätta de teoretiska momenten lika med stickprovsmomenten.

   Väntevärdet av binomialfördelningen ges av $M_1 = np$, där $n = 62$ i detta fall. Eftersom vi endast har en observation, $x = 55$, ges det motsvarande stickprovsmomentet av denna observation. För att skatta p sätter vi $M_1 = 55$ och löser ut p. Detta ger $p^* = 55/62$.

   (c) Låt X vara det totala antalet lyckade försök, som nu är $\mathrm{Bin}(50, p)$. Vi söker

$$
\begin{aligned}
\mathsf{P}(X \geq 45) &= \mathsf{P}(X = 45) + \mathsf{P}(X = 46) + \mathsf{P}(X = 47) + \mathsf{P}(X = 48) + \mathsf{P}(X = 49) + \mathsf{P}(X = 50) \\
&= \sum_{k=45}^{50} \binom{50}{k} p^k (1-p)^{50-5} \\
&= 0.1770259 + 0.1510408 + 0.1009024 + 0.04950245 + 0.0158601 + 0.002489895 \\
&= 0.4968216
\end{aligned}
$$

5. (a) Log-likelihoodfunktionen ges av

$$\ell(x) = \sum_{i=1}^{n} \log(x_i) - \log(a) - \frac{1}{2a} x_i^2$$

   där $n = 52$. Vi deriverar denna funktion med avseende på a

$$\frac{\partial\, \ell(x)}{\partial\, a} = -\frac{n}{a} + \frac{1}{2a^2} \sum_{i=1}^{n} x_i^2.$$

   Vi sätter derivatan till noll och löser ut a, vilket ger

$$a^*_{ML} = \frac{1}{2n} \sum_{i=1}^{n} x_i^2.$$

   Om vi sätter in datan får vi

$$a^*_{ML} = \frac{2848.5}{2 \cdot 52} = 27.39.$$

   (b) Låt X vara en Rayleighfördelad slumpvariabel med parameter $a = 27.39$. Vi vill beräkna

$$
\begin{aligned}
\mathsf{P}(3 \leq X \leq 25) &= \int_3^{25} \frac{x}{a} e^{-\frac{x^2}{2a}}\, dx = \left[-e^{-\frac{x^2}{2a}} \right]_3^{25} \\
&= e^{-\frac{3^2}{2a}} - e^{-\frac{25^2}{2a}} = 0.848
\end{aligned}
$$

   (c) Notera att $\mathsf{E}(X^2) = \mathsf{V}(X) + \mathsf{E}(X)^2$. Vi har att

$$
\begin{aligned}
\mathsf{E}(a^*_{ML}) &= \mathsf{E}\left(\frac{1}{2n} \sum_{i=1}^{n} X_i^2 \right) = \frac{1}{2n} \sum_{i=1}^{n} \mathsf{E}(X_i^2) \\
&= \frac{1}{2n} \sum_{i=1}^{n} \mathsf{V}(X_i) + \mathsf{E}(X_i)^2 = \frac{1}{2n} \sum_{i=1}^{n} (2 - \pi/2)a + a\pi/2 \\
&= \frac{1}{2n} \sum_{i=1}^{n} 2a = \frac{1}{2n} n 2a = a.
\end{aligned}
$$

   Alltså är skattaren väntevärdesriktig.

(d) Vi börjar med att beräkna skattningens varians

$$\mathsf{V}(a^*_{ML}) = \mathsf{V}\left(\frac{1}{2n}\sum_{i=1}^{n} X_i^2\right) = \frac{1}{4n^2}\sum_{i=1}^{n}\mathsf{V}(X_i^2) = \frac{1}{4n^2}n4a^2 = a^2/n.$$

Via normalapproximation får vi ett konfidensintervall som

$$I_a = \left[a^*_{ML} \pm z_{\alpha/2}\frac{a^*_{ML}}{\sqrt{n}}\right] = \left[27.39 \pm 1.96 \cdot \frac{27.39}{\sqrt{52}}\right] = [19.94,\ 34.83].$$

6. (a) För kornsort i låt T_i vara vikten för skörden från det torkade kornet och I_i vikten för skörden för den icke-torkade. Vi använder stickprov i par eftersom det antagligen finns en stor variation mellan olika kornsorter, och behöver då att skillnaderna $D_i = T_i - I_i$ kan antas vara oberoende och normalfördelade $\mathsf{N}(\Delta, \sigma^2)$. Vi vill testa $H_0 : \Delta = 0$ mot $H_1 : \Delta > 0$.

(b) Vi börjar med att beräkna skillnaderna D_i:

År	1899							1900			
Kornsort	1	2	3	4	5	6	7	8	9	10	11
D_i	106	-20	101	-33	72	-36	62	38	-70	127	24

Från dessa kan vi skatta medelvärdet $\bar{D} = 33.72727$ och standardavvikelsen $s = 66.17113$. Vi beräknar teststorheten

$$T = \frac{33.72727}{66.17113/\sqrt{11}} = 1.690476.$$

Eftersom vi har ett ensidigt test jämför vi denna med $t_{0.1}(10) = 1.3722$. Eftersom $T > t_{0.1}(10)$ kan vi förkasta nollhypotesen.

(c) Vi antar att skillnaderna från 1899 är $\mathsf{N}(\Delta_1, \sigma_1^2)$, och skillnaderna 1900 är $\mathsf{N}(\Delta_2, \sigma_2^2)$. För att undersöka om vi kan göra ett poolat test måste vi först testa $H_0 : \sigma_1 = \sigma_2$ mot $H_1 : \sigma_1 \neq \sigma_2$. Notera att om vi kan förkasta nollhypotesen här så vore antagandena för testet i (b) inte uppfyllda. Vi börjar med att skatta varianserna för de två åren, och får $s_1^2 = 4029.667$, $s_2^2 = 6502.917$. Vi beräknar teststorheten

$$T = \frac{s_2^2}{s_1^2} = 1.61376.$$

Vi utför testet på signifikansnivå 95% och kan då förkasta nollhypotesen om $T > f_{0.05}(3, 6)$. Från tabell får vi att $f_{0.05}(3, 6) = 4.75$ och kan alltså inte förkasta H_0.

(d) Från (c) vet vi att vi kan anta att $\sigma_1 = \sigma_2$ och vill nu testa $H_0 : \Delta_1 = \Delta_2$ mot $H_1 : \Delta_1 \neq \Delta_2$. Notera igen att om vi kan förkasta H_0 så skulle antagandena för testet i (b) inte vara uppfyllda. Vi beräknar $\bar{D}_1 = 36$ och $\bar{D}_2 = 29.75$. Vi behöver också den poolade skattningen av den gemensamma variansen, $s_p^2 = \frac{6}{9}s_1^2 + \frac{3}{9}s_2^2 = 4854.083$. Ett konfidensintervall för skillnaden $\Delta_1 - \Delta_2$ ges nu av

$$\left[\bar{D}_1 - \bar{D}_2 \pm t_{0.05}(9)s_p\sqrt{\frac{1}{4} + \frac{1}{7}}\right] = [-73.80,\ 86.30].$$

Eftersom detta täcker noll kan vi inte förkasta H_0.