1. (a) Enligt definitionen av betingad sannolikhet är $\mathsf{P}(A \cap B) = \mathsf{P}(B|A)\mathsf{P}(A) = 0.1 \cdot 0.3 = 0.03$. Vi använder nu additionssatsen för två händelser och får $\mathsf{P}(B) = \mathsf{P}(A \cup B) + \mathsf{P}(A \cap B) - \mathsf{P}(A) = 0.5 + 0.03 - 0.3 = 0.23$.

   (b) Vi har $\mathsf{E}(Y) = \mathsf{E}(X+Y) - \mathsf{E}(X) = 2$ och eftersom $\mathsf{V}(X+Y) = \mathsf{V}(X) + \mathsf{V}(Y) + 2\mathsf{C}(X,Y)$ så är $\mathsf{V}(Y) = \mathsf{V}(X+Y) - \mathsf{V}(X) - 2\mathsf{C}(X,Y) = 4 - 2 - 1 = 1$. Alltså är $Y \sim \mathsf{N}(2,1)$ och vi har
$$\mathsf{P}(Y > 1) = 1 - \Phi\left(\frac{1-2}{1}\right) = 1 - \Phi(-1) = 0.8413.$$

2. (a) Vi har
$$\int_0^{10}\int_0^5 (10-x)(5-y)dydx = \int_0^{10}(10-x)\left[5y - \frac{y^2}{2}\right]_0^5 dx$$
$$= \frac{25}{2}\int_0^{10}(10-x)dx = \frac{25}{2}\left[10x - \frac{x^2}{2}\right]_0^{10} = 625.$$
   Alltså är $C = 625$.

   (b) Vi har $f_{X|y}(x) = f(x,y)/f_Y(y)$ och vi måste alltså först beräkna $f_Y(y)$:
$$f_Y(y) = \frac{1}{625}\int_0^{10}(10-x)(5-y)dx = \frac{5-y}{625}\left[10x - \frac{x^2}{2}\right]_0^{10} = \frac{(5-y)50}{625} = \frac{2(5-y)}{25}.$$
   Alltså är
$$f_{X|y}(x) = \frac{f(x,y)}{f_Y(y)} = \frac{(10-x)(5-y)}{625}\frac{25}{2(5-y)} = \frac{(10-x)}{50}.$$

   (c) X och Y är oberoende om $f_{X|y}(x) = f_X(x)$. Vi har
$$f_X(x) = \frac{1}{625}\int_0^5 (10-x)(5-y)dy = \frac{10-x}{625}\left[5y - \frac{y^2}{2}\right]_0^5 = \frac{(10-x)}{,}625\frac{25}{2} = \frac{10-x}{50}$$
   och alltså är X och Y oberoende.

3. (a) Vi antar att temperaturen är normalfördelad, $T_i \sim \mathsf{N}(\mu, \sigma^2)$, och vill testa $H_0 : \mu = 20$ mot $H_1 : \mu < 20$. Vi gör detta på signifikansnivå $\alpha = 0.05$. Vi har stickprovsmedelvärde $\bar{t} = 19.7250$ och stickprovsvarians $s^2 = 0.5166$. Vi beräknar ett ensidigt konfidensintervall för μ enligt:
$$I_\mu = \left(-\infty, \bar{t} + t_\alpha(9)\frac{s}{\sqrt{10}}\right) = (-\infty, 20.1417).$$
   Eftersom 20 ligger i intervallet så kan vi inte förkasta H_0.

   (b) Med $T \sim \mathsf{N}(21.5, 2^2)$ har vi
$$\mathsf{P}(18 \le T \le 26) = \Phi\left(\frac{26-21.5}{2}\right) - \Phi\left(\frac{18-21.5}{2}\right) = \Phi(2.25) - \Phi(-1.75)$$
$$= \Phi(2.25) - (1 - \Phi(1.75)) = 0.9878 - (1 - .9599) = 0.9477.$$
   Den sökta sannolikheten är alltså $1 - 0.9477 = 0.0523$.

4. (a) Låt M vara medianen i fördelningen, vi vill testa $H_0 : M = 20$ mot $H_1 : M < 20$. Vi gör detta på signifikansnivå $\alpha = 0.05$. Vi har $Q_+ = \{$ antal värden över $M\} = 3$. Om H_0 är sann är $Q_+ \sim$Bin$(10, 0.5)$ och vi beräknar p-värdet för testet som

$$\begin{aligned}
p &= \mathsf{P}(Q_+ \leq 3 | H_0) = p(0) + p(1) + p(2) + p(3) \\
&= \binom{10}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{10} + \binom{10}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{9} + \binom{10}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{8} + \binom{10}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{7} \\
&= \frac{1}{2^{10}}\left(\binom{10}{0} + \binom{10}{1} + \binom{10}{2} + \binom{10}{3}\right) = \frac{1}{2^{10}}(1 + 10 + 45 + 120) = \frac{176}{1024} = 0.1719.
\end{aligned}$$

Eftersom $p > 0.05$ kan vi inte förkasta H_0.

(b) Vi bör använda ett teckentest eller något annat icke-parametriskt test ifall antagandet om normalfördelning inte stämmer.

(c) Testest styrka ges av sannolikheten att inte utföra ett fel av typ 2, dvs styrkan är $1 - \beta$ där $\beta = \mathsf{P}(\text{Förkasta ej } H_0 | H_1 \text{ sann})$.

5. (a) Det är rimligt att anta att födelsedagarna hos de olika spelarna är oberoende av varandra, och vi har därför att $X \sim$Bin(n, p), där $n = 23$ och $p = 0.5$ om relativ ålderseffekt inte finns.

(b) Bland spelarna är 15 personer födda under de sex första månaderna. Eftersom vi i detta fall har en observation $x = 15$ av en binomialfördelning ges log-likelihood funktionen av

$$l(p) = \log f(x) = \log\left(\binom{n}{x} p^x (1-p)^{n-x}\right) = \log(\binom{n}{x} + x\log(p) + (n-x)\log(1-p)$$

Vi deriverar $l(p)$ med avseende på p och sätter derivatan till noll:

$$0 = \frac{x}{p} - \frac{n-x}{1-p} \Rightarrow \frac{n-x}{1-p} = \frac{x}{p} \Rightarrow p(n-x) = x(1-p) \Rightarrow pn = x \Rightarrow p = \frac{x}{n}$$

Vi sätter in $x = 15, n = 23$ och får $p^* = 15/23 = 0.6522$.

(c) Vi har

$$\begin{aligned}
\mathsf{E}(p^*) &= \mathsf{E}\left(\frac{x}{n}\right) = \frac{\mathsf{E}(x)}{n} = \frac{np}{n} = p \\
\mathsf{V}(p^*) &= \mathsf{V}\left(\frac{x}{n}\right) = \frac{\mathsf{V}(x)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}
\end{aligned}$$

(d) Vi vill testa $H_0 : p = 0.5$ mot $H_1 : p > 0.5$. Vi använder normalapproximation och beräknar teststorheten $T = (p^* - p_0)/\sqrt{p_0(1-p_0)/n} = (15/23 - 0.5)/\sqrt{0.5^2/23} = 1.4596$. p-värdet för testet blir $1 - \Phi(1.4596) = 0.0722$. Eftersom $p < \alpha$ kan vi förkasta H_0. Det verkar finnas en relativ ålderseffekt. Om vi använder direktmetoden och beräknar p-värdet för testet utan att använda normalapproximation får vi 0.0466 och kan då alltså också förkasta H_0 med $\alpha = 0.05$.

6. (a) Vi har ett faktorförsök med tvåsidig indelning, och ansätter modellen

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

där $\varepsilon_{ijk} \sim \mathsf{N}(0, \sigma^2)$, $i = 1$ om Faktor A har låg nivå, och $i = 2$ annars, $j = 1$ om Faktor B har låg nivå och $j = 2$ annars. Slutligen är $k = 1, 2, 3$.

(b) Vi börjar med att beräkna de olika medelvärdena som behövs för att beräkna tabellen: $\bar{y}_{11} = 1.9233\bar{y}_{12} = 5.1567, \bar{y}_{21} = 3.8067, \bar{y}_{22} = 8.1800, \bar{y} = 4.7667$, samt

$$\begin{aligned}
\bar{y}_{\cdot 1} &= (\bar{y}_{11} + \bar{y}_{21})/2 = 2.8650, \\
\bar{y}_{\cdot 2} &= (\bar{y}_{12} + \bar{y}_{22})/2 = 6.6683, \\
\bar{y}_{1\cdot} &= (\bar{y}_{11} + \bar{y}_{12})/2 = 3.5400, \\
\bar{y}_{2\cdot} &= (\bar{y}_{21} + \bar{y}_{22})/2 = 4.4817.
\end{aligned}$$

Vi kan använda dessa för att beräkna följande kvadratsummor

$$\begin{aligned}
SS_A &= 6\left((\bar{y}_{1\cdot} - \bar{y})^2 + (\bar{y}_{2\cdot} - \bar{y})^2\right) = 18.056, \\
SS_B &= 6\left((\bar{y}_{\cdot 1} - \bar{y})^2 + (\bar{y}_{\cdot 2} - \bar{y})^2\right) = 43.3960, \\
SS_{Tot} &= \sum_{ijk}(y_{ijk} - \bar{y})^2 = \sum_{ijk} y_{ijk}^2 - 12\bar{y}^2 = 66.8353.
\end{aligned}$$

Vi behöver också SS_E och för att beräkna den beräknar vi först stickprovsvariansen för vardera av de fyra faktornivåerna: $s_{11}^2 = 1.1874, s_{12}^2 = 0.0729, s_{21}^2 = 0.1369, s_{22}^2 = 0.8067$. Vi kan nu beräkna

$$SS_E = 2(s_{11}^2 + s_{12}^2 + s_{21}^2 + s_{22}^2) = 4.4080.$$

Slutligen är $SS_{AB} = SS_{Tot} - SS_E - SS_A - SS_B = 0.9747$.

(c) Baserat på kvadratsummorna ovan kan vi nu fylla i alla fält i variansanalystabellen och får

Variation	Kvadratsumma	Frihetsgrader	Medelkvadrat	Teststorhet
Faktor A	18.056	1	9.5156	32.770
Faktor B	43.396	1	43.396	78.758
Faktor AB	0.9747	1	9.5156	1.7690
Residual	4.4080	8	0.5510	
Total	66.8353	11		

(d) Det kritiska värdet för teststorheten är $F_{0.05}(1, 8) = 5.31$. Eftersom $F = 1.769 < 5.31$ kan vi inte förkasta H_0, och vi har alltså ingen samspelseffekt.

(e) Det kritiska värdet för teststorheterna är också här $F_{0.05}(1, 8) = 5.31$. Eftersom teststorheterna för båda faktorerna (32.77 och 78.758) är större än 5.31 är båda effekterna signifikanta.