1. (a) Låt $Z = 2X + 5Y$, vi har då $\mathsf{E}(Z) = 2 \cdot 1 + 5 \cdot 2 = 12$ och $\mathsf{V}(Z) = 2^2 \cdot 4 + 5^2 \cdot 3 = 91$. Alltså, $Z \sim \mathsf{N}(12, 91)$ och

$$\mathsf{P}(Z > 4) = 1 - \Phi\left(\frac{4-12}{\sqrt{91}}\right) = \Phi(0.8386) \approx 0.799$$

(b) Vi har $\mathsf{P}(A^c \cap B^c) = \mathsf{P}((A \cup B)^c) = 1 - \mathsf{P}(A \cup B)$, och

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B) = 0.5 + 0.3 - 0.1 = 0.7.$$

Alltså är $\mathsf{P}(A^c \cap B^c) = 1 - 0.7 = 0.3$.

2. (a) Vi har att

$$f_{X,Y}(x,y) = f_X(x) f_{Y|x}(y) = \begin{cases} e^{-x} & \text{om } x \geq 0 \text{ och } 0 < y < x \\ 0 & \text{för övrigt} \end{cases}$$

Alltså ges $f_Y(y)$ av

$$f_Y(y) = \int f_{X,Y}(x,y)dx = \int_y^\infty e^{-x} dx = e^{-y}$$

Så vi har $Y \sim$Exp(1).

(b) Eftersom $Y \sim$Exp(1) har vi $\mathsf{E}(Y) = \mathsf{V}(Y) = 1$. Vi ser att $\mathsf{E}(X) = \mathsf{E}(Y^2)$ och får därför $\mathsf{E}(X) = \mathsf{E}(Y^2) = \mathsf{V}(Y) + \mathsf{E}(Y)^2 = 1 + 1 = 2$. Alternativt kan vi använda partialintegrering och få $\mathsf{E}(X)$ som

$$\mathsf{E}(X) = \int_0^\infty x^2 e^{-x} dx = [-x^2 e^{-x}]_0^\infty + 2\int x e^{-x} dx = 0 + 2 \cdot 1 = 2$$

3. (a) Låt X vara antal felaktiga kostymer, vi har då $X \sim$Bin$(15, 0.05)$ och vi söker

$$\begin{aligned} \mathsf{P}(X > 1) &= 1 - \mathsf{P}(X \leq 1) = 1 - p_X(0) - p_X(1) \\ &= 1 - \binom{15}{0} 0.05^0 \cdot 0.95^{15} - \binom{15}{1} 0.05^1 \cdot 0.95^{14} \\ &= 1 - 0.95^{15} - 15 \cdot 0.05 \cdot 0.95^{14} \\ &= 0.171 \end{aligned}$$

(b) Låt $X \sim$Bin$(100, p)$. Vi vill utföra hypotestestet

$$\begin{aligned} H_0 &: p = 0.05 \\ H_1 &: p > 0.05 \end{aligned}$$

Vi väljer signifikansnivå $\alpha = 0.05$. Med direktmetoden beräknar vi p-värdet för testet som

$$\mathsf{P}(X \geq 7) = 1 - \mathsf{P}(X \leq 6) = 1 - \sum_{k=0}^{6} \binom{100}{k} 0.05^k \cdot 0.95^{15-k} = 0.2340$$

Eftersom $0.2340 > \alpha$ kan vi inte förkasta H_0. Kontrollen påvisar alltså inte att felandelen i produktionen är över 5%.

4. (a) Det är ett försök med så kallad ensidig indelning. Låt faktor A vara kolesterolhalten, med faktornivåer α_i, vi har då modellen

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij},$$

där $\varepsilon_{ij} \sim \mathsf{N}(0, \sigma^2)$.

(b) Vi har

$$\bar{y} = \frac{1}{4}(3.4756 + 2.8228 + 2.4778 + 3.0077) = 2.9460$$

och $Q_A = \sum_{ij}(\bar{y}_i - \bar{y})^2 = 50\sum_i(\bar{y}_i - \bar{y})^2 = 25.9344$. Variansanalystabellen blir

Variation	Kvadratsumma	Frihetsgrader	Medelkvadratsumma	Teststorhet
Faktor A	25.934	3	8.6448	9.2517
Residual	183.143	196	0.9344	
Total	209.0779	199		

(c) Vi vill utföra testet

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$$
$$H_1 : \text{åtminstone ett } \alpha_i \neq 0$$

Teststorheten är $F(3, 196)$-fördelad. Enligt formelsamlingen har vi

$$F_{0.05}(3, 196) \approx F_{0.05}(3, \infty) = 2.6507.$$

Det kritiska området för testet är $C_\alpha = \{T : T > 2.6507\}$. Eftersom $T_{obs} = 9.2517$ ligger i C_α kan vi förkasta nollhypotesen.

5. (a) Parametrarna skattas som

$$\boldsymbol{\beta}^* = \begin{pmatrix} \beta_0^* \\ \beta_1^* \\ \beta_2^* \end{pmatrix} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} = \begin{pmatrix} 0.0259 \\ -0.0037 \\ 0.0552 \end{pmatrix}$$

(b) Vi har $Q_0 = \mathbf{Y}^T\mathbf{Y} - \boldsymbol{\beta}^{*T}\mathbf{X}^T\mathbf{Y} = 0.7460$ och $s = \sqrt{Q_0/(15-3)} = 0.2493$. Med $n = 15$ observationer och $p + 1 = 3$ parametrar är antalet frihetsgrader i skattningen alltså 12.

(c) Konfidensintervallen ges av

$$I_{\beta_1} = \left(\beta_1^* \pm t_{\alpha/2}(12)s\sqrt{((\mathbf{X}^T\mathbf{X})^{-1})_{2,2}}\right) = (-0.0248,\ 0.0173)$$
$$I_{\beta_2} = \left(\beta_2^* \pm t_{\alpha/2}(12)s\sqrt{((\mathbf{X}^T\mathbf{X})^{-1})_{3,3}}\right) = (0.0364,\ 0.0740)$$

(d) Låt $\mathbf{x}_0 = [1,\ 37,\ 10]^T$. Vi har då $\mu_Y^*(\mathbf{x}_0) = \boldsymbol{\beta}^{*T}\mathbf{x}_0 = 0.4401$ och $\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 = 0.1471$. Intervallet ges av

$$I_{\mu_Y(\mathbf{x}_0)} = \left(\mu_Y^*(\mathbf{x}_0) \pm t_{\alpha/2}(n-p-1)s\sqrt{\mathbf{x}_0^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0}\right) = (0.2317,\ 0.6484)$$

(e) Ja, eftersom intervallet för β_1 täcker noll är denna parameter inte signifikant. Alltså kan vi inte påvisa att gastemperaturen har en signifikant påverkar på absorptionen och vi skulle kunna ansätta den enklare modellen

$$y_i = \beta_0 + \beta_2 x_{2i} + \varepsilon_i$$

6. (a) Låt x_F och x_E beteckna mätningarna hos icke-rökarna för och efter studien. På motsvarande sätt låt y_F och y_E beteckna mätningarna hos rökarna. De olika skattningarna fås som stickprovsmedelvärde och stickprovsvarians hos datan för respektive grupp. Resultaten sammanfattas i följande tabell:

	Icke-rökare		Rökare	
	Väntevärde	Varians	Väntevärde	Varians
Före	132.8333	4.5667	134.2000	5.2000
Efter	131.8333	12.9667	130.2000	8.7000

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

$$H_0 : \sigma_{I_F} = \sigma_{R_F}$$
$$H_1 : \sigma_{I_F} \neq \sigma_{R_F}$$

på nivå $\alpha = 0.05$. Vi beräknar teststorheten $T_{obs} = s^2_{I_F}/s^2_{I_F} = 5.2/4.5667 = 1.1387$, 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.

(c) Vi har samma modell som i (a) men antar att $\sigma_{I_F} = \sigma_{R_F}$ och vill nu testa

$$H_0 : \mu_{I_F} = \mu_{R_F}$$
$$H_1 : \mu_{I_F} \neq \mu_{R_F}$$

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

$$s_p^2 = \frac{4s^2_{RF} + 5s^2_{IF}}{9} = \frac{43.6333}{9} = 4.8481$$

Vi beräknar teststorheten

$$T_{obs} = \frac{\bar{x}_F - \bar{y}_F}{s_p\sqrt{\frac{1}{4} + \frac{1}{5}}} = -0.9253$$

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.

(d) Vi beräknar skillnaderna i blodtryck för varje rökare:

Person nr	7	8	9	10	11
Skillnad	-5	-2	-2	-6	-5

Vi ansätter att skillnaden är $\mathsf{N}(\Delta, \sigma^2)$-fördelad och vi vill nu testa

$$H_0 : \Delta = 0$$
$$H_1 : \Delta < 0$$

på nivå $\alpha = 0.05$. Vi har $\bar{D} = -4$ och $s^2_D = 14/4 = 3.5$. Teststorheten är nu

$$T_{obs} = \frac{\bar{D}}{s_D/\sqrt{5}} = -4.7809$$

Under H_0 är teststorheten $t(4)$-fördelad. Eftersom $T_{obs} < -t_\alpha(4) = -2.1318$ kan vi förkasta H_0.