1. (a) Om T är tiden så söker vi

$$\mathsf{P}(T > 2) = \int_2^\infty f(t)dt = \int_2^\infty \frac{1}{5}e^{-t/5}dt = [-e^{-t/5}]_2^\infty = e^{-2/5}.$$

(b) Den sökta sannolikheten fås som $\mathsf{P}(\text{Alla tre kommer}) + \mathsf{P}(\text{Två av tre kommer})$. Låt A vara händelsen att Anna kommer, B vara händelsen att Bert kommer och C vara händelsen att Carl kommer. Vi har då

$$\begin{aligned}
\mathsf{P}(\text{Alla tre kommer}) &= \mathsf{P}(A \cap B \cap C) = 0.8 \cdot 0.6 \cdot 0.9 = 0.432 \\
\mathsf{P}(\text{Två av tre kommer}) &= \mathsf{P}(A \cap B \cap C^c) + \mathsf{P}(A \cap B^c \cap C) + \mathsf{P}(A^c \cap B \cap C) \\
&= 0.8 \cdot 0.6 \cdot 0.1 + 0.8 \cdot 0.4 \cdot 0.9 + 0.2 \cdot 0.6 \cdot 0.9 \\
&= 0.108 + 0.288 + 0.048 = 0.444
\end{aligned}$$

Alltså ges den sökta sannolikheten av $0.432 + 0.444 = 0.876$.

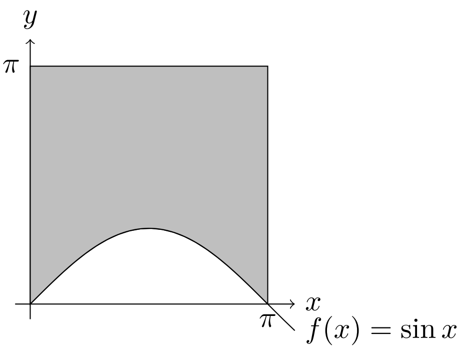
2. (a) Vi har

$$\mathsf{E}(\sin(X)) = \int_{-\infty}^\infty \sin(x)f(x)dx = \int_0^\pi \sin(x)\frac{1}{\pi}dx = \frac{1}{\pi}(-\cos(\pi) + \cos(0)) = \frac{2}{\pi}$$

(b) Eftersom X och Y är oberoende ges den simultana täthetsfunktionen för (X, Y) av produkten av marginalfördelningarna:

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \frac{1}{\pi^2} & \text{Om } 0 \le x \le \pi \text{ och } 0 \le y \le \pi \\ 0 & \text{för övrigt} \end{cases}$$

(c) Den sökta sannolikheten ges av arean av det skuggade området delat med π^2:



$$\begin{aligned}
\mathsf{P}(Y > \sin(X)) &= \int_0^\pi \int_{\sin(x)}^\pi \frac{1}{\pi^2}dydx = \frac{1}{\pi^2}\int_0^\pi [y]_{\sin(x)}^\pi dx = \frac{1}{\pi^2}\int_0^\pi \pi - \sin(x)dx \\
&= \frac{1}{\pi^2}[\pi x + \cos(x)]_0^\pi = \frac{1}{\pi^2}(\pi^2 - 2) = 1 - \frac{2}{\pi^2} \approx 0.7974
\end{aligned}$$

3. (a) Sannolikheten ges av

$$\mathsf{P}(L > 20) = \mathsf{P}\left(\frac{L - 22.5}{5} > \frac{20 - 22.5}{5}\right) = 1 - \Phi\left(\frac{20 - 22.5}{5}\right) = \Phi(0.5) = 0.6915$$

(b) Låt $p = \mathsf{P}(L > 20) = 0.6915$ och låt X vara antalet fiskar som är minst 20cm långa. Vi har att X ∼Bin(5,p) och söker sannolikheten

$$\mathsf{P}(X \geq 3) = \sum_{k=3}^{5} \mathsf{P}(X = k) = \sum_{k=3}^{5} \binom{5}{k} p^k (1-p)^{5-k} = 0.3147 + 0.3527 + 0.1581 = 0.8254$$

4. Låt X_i beteckna mätningarna i Nordstan och Y_i mätningarna i Haga. Antag att mätningarna är oberoende samt att $X_i \sim \mathsf{N}(\mu_x, \sigma^2)$ och $Y_i \sim \mathsf{N}(\mu_y, \sigma^2)$. Vi vill testa $H_0 : \mu_x = \mu_y$ mot $H_1 : \mu_x \neq \mu_y$. Ett konfidensintervall för $\mu_x - \mu_y$ ges av

$$I_{\mu_x - \mu_y} = \left[\bar{x} - \bar{y} \pm t_{\alpha/2}(12) s_p \sqrt{\frac{2}{7}} \right],$$

där $\bar{x} = 8.8429, \bar{y} = 11.3000$ är stickprovsmedelvärdena och $s_p = \sqrt{(s_x^2 + s_y^2)/2} = 2.1122$ är en poolad skattning av σ, där s_x^2 och s_y^2 är stickprovsvarianserna för X respektive Y. Med $\alpha = 0.1$ får vi $I_{\mu_x - \mu_y} = [0.4449, 4.4694]$. Sätter vi däremot $\alpha = 0.05$ får vi $I_{\mu_x - \mu_y} = [-0.0028, 4.9171]$. Vi kan alltså förkasta H_0 på nivå $\alpha = 0.1$ men inte på nivå $\alpha = 0.05$.

5. (a) Vi vill skatta parametrarna i regressionsmodellen $Y = \beta_0 + \beta_1 x + \varepsilon$, där Y är fångsten, x är året och $\varepsilon \sim \mathsf{N}(0, \sigma^2)$. Vi har $n = 14$ observationer och beräknar

$$\bar{x} = \frac{1999 + \ldots + 2012}{14} = 2005.5, \qquad \bar{y} = \frac{3678 + \ldots + 6268}{14} = 5104.786$$

$$S_{xx} = \sum_{i=1}^{14} (x_i - \bar{x})^2 = 227.5, \qquad S_{yy} = \sum_{i=1}^{14} (y_i - \bar{y})^2 = 23055896$$

$$S_{xy} = \sum_{i=1}^{14} (y_i - \bar{y})(x_i - \bar{x}) = 42026.5$$

Skattningarna av parametrarna ges nu av $\beta_1^* = \frac{S_{xy}}{S_{xx}} = 184.7319$, $\beta_0^* = \bar{y} - \beta_1^* \bar{x} = -365375$ och $s^2 = \frac{syy - \beta_1 \bar{y}}{n-2} = \frac{15292263}{n-2} = 1274355$.

(b) Med $\alpha = 0.1$ och $t_{\alpha/2}(n-2) = -1.782288$ får vi intervallet

$$I_{\beta_1} = \left[\beta_1^* \pm t_{\alpha/2}(n-2) \sqrt{\frac{s^2}{S_{xx}}} \right] = [184.7319 \pm 1.78 \cdot 74.84] = [51.34, 318.12]$$

Eftersom konfidensintervallet inte inkluderar 0 så verkar det som att antal kg uppfiskade humrar förändras med tiden.

(c) Skattningen av det förväntade antalet kilo humrar som kommer fångas år 2015 ges av $\hat{\mu}_{Y|x}(2015) = \beta_0 + \beta_1 2015 = 6859.7$. Med $\alpha = 0.05$ och $t_{\alpha/2}(n-2) = -2.178813$ fås ett 95% prediktionsintervall för kvantiteten som

$$\begin{aligned} I_{Y|x(2012)} &= \left[\hat{\mu}_{Y|x}(2015) \pm t_{\alpha/2}(n-2) s \sqrt{1 + \frac{1}{n} + \frac{(2015 - \bar{x})^2}{S_{xx}}} \right] \\ &= [6859.7 \pm 2.18 \cdot 1367.8] = [3879.5, 9839.95] \end{aligned}$$

Givet att den ansatta modellen stämmer kommer alltså med 95% säkerhet mellan 3879.5 kg och 9839.95 kilo hummer fiskas upp under 2015.

6. (a) Vi vill testa

$$H_0 : M \geq 3 \tag{1}$$
$$H_1 : M < 3 \tag{2}$$

där M är medianen i fördelningen. Beräkna differenserna $D_i = y_i - 3$ för de uppmätta värdena y_i och bilda teststorheten

$$Q_+ = \sum_{i=1}^{12} \mathbb{I}(D_i \geq 0) = 4 \tag{3}$$

Under nollhypotesen så skall sannolikheten att få ett positivt värde på D_i vara lika stor som att få ett negativt och därför måste Q_+ vara binomialfördelad, $Q_+ \sim \text{Bin}(14{,}0.5)$. Vi använder detta för att beräkna (eller slå upp i tabell)

$$p = \mathsf{P}(Q_+ \leq 4) = 0.1208 \tag{4}$$

Eftersom $p > 0.01$ så vi kan inte förkasta nollhypotesen med 1% signifikans. Alltså kan man inte med tillräckligt stor säkerhet påvisa att bearbetningsanläggningen uppfyller miljökraven.

(b) Skall vi använda oss av Wilcoxon ranktest måste vi anta att fördelningen är symmetrisk, annars är testet inte giltigt. Fördelen med testet är att det i regel har högre styrka än teckentestet, eftersom det använder mer information än enbart tecknen på D_i.

(c) För att vi ska få använda normal-baserade test måste datan antingen vara normalfördelad eller så måste vi ha tillräckligt många observationer så att centrala gränsvärdessatsen kan användas för att motivera användandet av testet. Om vi får använda ett normal-baserat test är fördelen att det har högre styrka än icke-parametriska test.