Problem 1

By definition of marginal probability,

$$P(X) = \sum_Y P(X, Y) = \sum_Y P(X|Y) P(Y) \quad ①$$

$$P(Y) = \sum_X P(X, Y) = \sum_X P(Y|X) P(X) \quad ②$$

And we know, $P(X, Y) = f(X) g(Y)$

Thus, from ①, $\sum_Y f(x) g(Y) = \sum_Y P(X|Y) P(Y)$

from ②, $\sum_X f(x) g(Y) = \sum_X P(Y|x) P(x)$

$$f(x) = P(x|Y) = P(X)$$

$$g(Y) = P(Y) = P(Y|X)$$

$$P(X) = P(X|Y), \quad P(Y) = P(Y|X)$$

Therefore, X and Y are independent.

$$P(x) P(Y) = \sum_y P(X, y) \sum_x P(x, Y)$$

$$= \sum_y f(X) g(y) \sum_x f(x) g(Y)$$

$$= f(X) g(Y) \sum_y g(y) \sum_x f(x) = f(X) g(Y) \sum_{x,y} f(x) g(y)$$

$$= f(X) g(Y) \sum_{x,y} P(x, y) = f(X) g(Y) = P(X, Y)$$

Problem 2

$$\mathbb{E}_{x\sim p(x)}[af(x)+bg(x)]$$

$$= \sum_x p(x)[af(x)+bg(x)]$$

$$= \sum_x [ap(x)f(x) + bp(x)g(x)]$$

$$= \sum_x ap(x)f(x) + \sum_x bp(x)g(x)$$

$$= a\sum_x p(x)f(x) + b\sum_x p(x)g(x)$$

$$= a\mathbb{E}_{x\sim p(x)}[f(x)] + b\mathbb{E}_{x\sim p(x)}[g(x)]$$

Thus, the expectation is a linear system.

Problem 3.

(a) likelihood: $\mathcal{L}(\theta=[a,b]; D=\{x^{(m)}\}_{m=1}^{m})$

$$= \prod_{m=1}^{m} \cancel{f}\,P(x^{(m)}; [a,b])$$

$$= \prod_{m=1}^{m} \frac{1}{b-a} = \frac{1}{(b-a)^m}$$

(b) maximum likelihood estimate of parameters:

$$\hat{\theta}_{MLE} = [\hat{a}, \hat{b}]_{MLE} = \underset{[a,b]}{\arg\max}\ \mathcal{L}(\theta)$$

$$= \underset{[a,b]}{\arg\max}\ L([a,b])$$

$$= \underset{[a,b]}{\arg\max}\ \log L([a,b])$$

$$= \underset{[a,b]}{\arg\max}\ \log\left[\prod_{m=1}^{m} f(x^{(m)}; [a,b])\right]$$

$$= \underset{[a,b]}{\arg\max}\ \log\left[\prod_{m=1}^{m} \frac{1}{b-a}\right]$$

$$= \underset{[a,b]}{\arg\max}\ \log\left[\frac{1}{(b-a)^m}\right]$$

$$= \underset{[a,b]}{\arg\max}\ (-m)\log(b-a)$$

$$\frac{\partial}{\partial a}\left(-m\log(b-a)\right) = -\left(\frac{-m}{b-a}\right) = \frac{m}{b-a}$$

We want to maximum the $\hat{\theta}_{MLE}$, therefore,

$$\hat{a}_{MLE} = \min(X_1, X_2, X_3, \cdots X_m)$$

$$\hat{b}_{MLE} = \max(X_1, X_2, X_3, \cdots X_m)$$

Problem 4

(c) For uniformly distributed,

$$\mathbb{E}(x) = \int_a^b \frac{x}{b-a}\,dx$$

$$= \frac{1}{b-a}\int_a^b x\,dx$$

$$= \frac{1}{b-a}\cdot\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a}\cdot\left(\frac{b^2}{2}-\frac{a^2}{2}\right)$$

$$= \frac{1}{2}\cdot\frac{b^2-a^2}{b-a}$$

$$= \frac{1}{2}\cdot(a+b) = \frac{a+b}{2}$$

$$\mathbb{E}[\hat{a}_{MLE}] = \frac{a+b}{2} \qquad \mathbb{E}[\hat{b}_{MLE}] = \frac{a+b}{2}$$

By Thm, $\mathbb{E}[\hat{\omega}_{MLE}] = \omega$

However, $\mathbb{E}[\hat{a}_{MLE}] = \frac{a+b}{2} \neq a = \min(x_1, x_2, x_3, \cdots x_m)$

$\mathbb{E}[\hat{b}_{MLE}] = \frac{a+b}{2} \neq b = \max(x_1, x_2, x_3, \cdots x_m)$

Thus, MLE is biased in this case.

(d) From (a), $\mathcal{L}(\theta = [a, b]; D = \{x^{(m)}\}_{m=1}^{m}) = \frac{1}{(b-a)^m}$

$m \to +\infty$, $(b-a)^m \to +\infty$, $\frac{1}{(b-a)^m} \to 0$

$\Longrightarrow$ $\mathcal{L}$ is a constant function.

$\Longrightarrow$ $x_1 = x_2 = x_3 = x_4 = \cdots x_m$

$\Longrightarrow$ $\frac{a+b}{2} = a = \min(x_1, x_2, x_3, \cdots x_m)$

$\frac{a+b}{2} = b = \max(x_1, x_2, x_3, \cdots x_m)$

Thus, MLE is asymptotically unbiased if $m \to +\infty$.