Ex 1.3.
Part a.

$$P(Y = a) = P(x = r, Y = a) + P(X = b, Y = a) + P(X = g, Y = a) = .34$$

Part b, directly:

$$P(X = g \mid Y = o) = \frac{P(X = g, Y = o)}{P(Y = o)} = \frac{.6 \cdot .3}{.36} = .5$$

Part b, using Bayes's theorem:

$$P(X = g \mid Y = o) = \frac{P(Y = o \mid X = g)P(X = g)}{P(Y = o)} = \frac{.3 \cdot .6}{.36} = .5$$

Ex 1.5. We're asked to show $\mathbb{E}[(f(x) - \mathbb{E}[f(x)])^2] = \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2$.

$$\begin{aligned}
&\mathbb{E}[(f(x) - \mathbb{E}[f(x)])^2] \\
&= \sum_x p(x)(f(x) - \mathbb{E}[f(x)])^2 \\
&= \sum_x p(x)(f(x)^2 - 2f(x)\mathbb{E}[f(x)] + \mathbb{E}[f(x)]^2) \\
&= \left(\sum_x p(x)f(x)^2\right) - \left(2\mathbb{E}[f(x)]\sum_x p(x)f(x)\right) + \left(\mathbb{E}[f(x)]^2 \sum_x p(x)\right) \\
&= \mathbb{E}[f(x)^2] - 2\mathbb{E}[f(x)]\mathbb{E}[f(x)] + \mathbb{E}[f(x)]^2 \\
&= \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2
\end{aligned}$$

Ex 1.6. We're asked to show that if X and Y are independent, then $\mathbb{E}_{X,Y}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

$$
\begin{aligned}
\mathbb{E}_{X,Y}[XY] &= \sum_x xy\ p(x,y) && \text{“joint” expectation} \\
&= \sum_x \sum_y xy\ p(x)\ p(y) && \text{since } X \text{ and } Y \text{ are independent} \\
&= \sum_x \sum_y (xp(x)\ yp(y)) && = \sum_x \left(xp(x) \sum_y \ yp(y) \right) \\
&= \sum_x (xp(x)\mathbb{E}[Y]) && = \mathbb{E}[Y] \sum_x xp(x) \\
&= \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}
$$

Gaussian distribution:

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Calculating the probability of data $\mathbf{x}$ for a given μ and σ^2:

$$p(\mathbf{x}\,|\mu,\sigma^2) = \prod_{i=1}^{N} \mathcal{N}(x_i|\mu,\sigma^2)$$

Calculating μ for maximum likelihood:

$$\mu_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Calculating σ^2 for maximum likelihood:

$$\sigma^2_{ML} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_{ML})^2$$