Stat 3011 Midterm 1 (Class Part)

Problem 1

All of this is summarized in the box on p. 25 in Wild and Seber.

On the first issue (what is the difference),

- In an **experiment**, the experimenter *determines* which experimental units receive which treatments (ideally by randomized assignment to treatment groups).
- In an **observational study**, which subject gets which treatment is outside the experimenter's control. Perhaps the subjects or their doctors choose. Perhaps it just happens (exposure to a toxic substance in the environment, for example).

On the second issue (what is the implication)

- A properly designed and executed **experiment** can reliably demonstrate *causation.*
- An **observational study** can only suggest possible causes. It *cannot* reliably establish *causation.* A properly designed and executed experiment is needed to follow up results implied by an observational study.

Problem 2

(a) The mean is

$$
\begin{aligned}
\sum_x x \operatorname{pr}(x) &= 1 \cdot \frac{1}{7} + 2 \cdot \frac{1}{7} + 3 \cdot \frac{1}{7} + 4 \cdot \frac{1}{7} + 5 \cdot \frac{1}{7} + 6 \cdot \frac{1}{7} + 7 \cdot \frac{1}{7} \\
&= \frac{1 + 2 + 3 + 4 + 5 + 6 + 7}{7} \\
&= \frac{28}{7} \\
&= 4
\end{aligned}
$$

(b) The standard deviation is

$$
\operatorname{sd}(X) = \sqrt{E\{(X - \mu)^2\}}
$$

where μ is the mean calculated in part (a), and

$$\begin{aligned}
E\{(X-\mu)^2\} &= \sum_x (x-\mu)^2 \operatorname{pr}(x) \\
&= (1-4)^2 \cdot \frac{1}{7} + (2-4)^2 \cdot \frac{1}{7} + (3-4)^2 \cdot \frac{1}{7} + (4-4)^2 \cdot \frac{1}{7} \\
&\qquad + (5-4)^2 \cdot \frac{1}{7} + (6-4)^2 \cdot \frac{1}{7} + (7-4)^2 \cdot \frac{1}{7} \\
&= \frac{(-3)^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2}{7} \\
&= \frac{9+4+1+0+1+4+9}{7} \\
&= \frac{28}{7} \\
&= 4
\end{aligned}$$

Thus $\operatorname{sd}(X) = \sqrt{4} = 2$.

Problem 3

(a) By the multiplication rule (using the assumed statistical independence)

$$\operatorname{pr}(\text{no accidents in 30 days}) = \operatorname{pr}(\text{no accident in one day})^3 0$$

So in order to answer this question we have to first answer the subsidiary question: what is the latter probability? By the complement rule

$$\begin{aligned}
\operatorname{pr}(\text{no accident in one day}) &= 1 - \operatorname{pr}(\text{an accident in one day}) \\
&= 1 - 0.002 \\
&= 0.998
\end{aligned}$$

Thus

$$\operatorname{pr}(\text{no accidents in 30 days}) = 0.998^3 0 = 0.941708$$

(b) The events in parts (a) and (b) of this problem are complementary, so by the complement rule each is one minus the other

$$\operatorname{pr}(\text{at least one accident in 30 days}) = 1 - 0.941708 = 0.058292$$

Problem 4

Curve A: skewed, long right tail, unimodal.
Curve B: symmetric, unimodal.
Curve C: skewed, long right tail, biimodal.
Curve D: symmetric, biimodal.