Stat 3011 Midterm 1 (Class Part)

Problem 1

These are described in Section 1.1.2 in Wild and Seber.

selection bias The sample is not representative of the population because of the method of selection (some form of systematic, non-random sampling).

nonresponse bias The nonresponders (individuals who do not return questionaires, who are not in when the interviewer calls, or who do not answer some questions) are different from responders.

self-selection (a special case of selection bias) individuals surveyed themselves choose to be in the sample (instant TV polls, internet polls, hotel and restaurant customer satisfacation forms).

question effects Wording of questions strongly influences answers.

survey-format effects Order of questions, length of survey, and general instructions applying to all questions can influence answers.

interviewer effects Interviewer tone of voice or facial expression can influence answers.

social desirability bias Some respondents are reluctant to give answers or admit to behavior that is generally considered socially undesirable.

Problem 2

(a) The mean is

$$
\begin{aligned}
\sum_x x \operatorname{pr}(x) &= 0 \cdot \frac{1}{9} + 1 \cdot \frac{2}{9} + 2 \cdot \frac{3}{9} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{1}{9} + \\
&= \frac{0 + 2 + 6 + 6 + 4}{9} \\
&= \frac{18}{9} \\
&= 2
\end{aligned}
$$

(b) The standard deviation is

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

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

$$\begin{aligned}
E\{(X-\mu)^2\} &= \sum_x (x-\mu)^2 \operatorname{pr}(x) \\
&= (0-2)^2 \cdot \frac{1}{9} + (1-2)^2 \cdot \frac{2}{9} + (2-2)^2 \cdot \frac{3}{9} + (3-2)^2 \cdot \frac{2}{9} \\
&\qquad + (4-2)^2 \cdot \frac{1}{9} \\
&= \frac{(-2)^2 \cdot 1 + (-1)^2 \cdot 2 + 0^2 \cdot 3 + 1^2 \cdot 2 + 2^2 \cdot 1}{9} \\
&= \frac{4+2+0+2+4}{9} \\
&= \frac{12}{9} \\
&= \frac{4}{3}
\end{aligned}$$

Thus

$$\operatorname{sd}(X) = \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}} = 1.154701.$$

Problem 3

You can't use the addition rule here because the events are not mutually exclusive (independent events aren't mutually exclusive). Thus some combination of the multiplication rule and complement rule must do the job.

The multiplication rule doesn't directly calculate events of the form "at least one" so we must use the complement rule first. The complementary event of "at least one" is "none," that is, the gambler loses all 10. And that is the kind of event that the multiplication rule handles

$$\begin{aligned}
\operatorname{pr}(\text{loses all 10}) &= \operatorname{pr}(\text{loses 1st}) \cdot \operatorname{pr}(\text{loses 2nd}) \cdots \operatorname{pr}(\text{loses 10th}) \\
&= \operatorname{pr}(\text{loses any one})^{10}
\end{aligned}$$

because all the bets have the same probability of a win or a loss.

By another application of the complement rule

$$\operatorname{pr}(\text{loses any one}) = 1 - \operatorname{pr}(\text{wins any one}) = 1 - 0.4736842 = 0.5263158$$

Hence

$$\operatorname{pr}(\text{loses any one})^{10} = 0.5263158^{10} = 0.001631038$$

and

$$\operatorname{pr}(\text{wins at least one}) = 1 - \operatorname{pr}(\text{loses all 10}) = 1 - 0.001631038 = 0.998369$$

Problem 4

(a)

$$E(X+Y) = E(X) + E(Y) = 3 + 3 = 6$$

(b)

$$E(X - Y) = E(X) - E(Y) = 3 - 3 = 0$$

(c)

$$\text{sd}(X + Y) = \sqrt{\text{sd}(X)^2 + \text{sd}(Y)^2} = \sqrt{1^2 + 1^2} = \sqrt{2}$$

(d)

$$\text{sd}(X - Y) = \sqrt{\text{sd}(X)^2 + \text{sd}(Y)^2} = \sqrt{1^2 + 1^2} = \sqrt{2}$$

Note: same as part (c) because $\text{sd}(X + Y) = \text{sd}(X - Y)$.

(e)

$$E(4X + 5) = 4E(X) + 5 = 4 \cdot 3 + 5 = 17$$

(f)

$$\text{sd}(4X + 5) = 4\,\text{sd}(X) = 4 \cdot 1 = 4$$