Nate Warner | STAT 300 | Summer 2024

PSET 5 - Due: Wednesday, July 10

1. An individual who has automobile insurance from a certain company is randomly selected. Let $X =$ the number of moving violations for which the individual was cited during the last 3 years. The probability mass function of X is given below.

x	0	1	2	3	4
$p(x)$	0.50	0.20	0.15	0.10	0.05

(a) Find the mean number of moving violations.

(b) Find the variance among the values of X.

(c) If an individual has X moving violations their insurance premium will change by $50X-20$. Find the mean amount of change to their premium, i.e., find $E(50X-20)$.

Remark. Let X be a discrete rv with set of possible values D and pmf $p(x)$. The **expected value** or **mean value** of X, denoted by $E(X)$ or μ_X or just μ, is

$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)$$

Let X have pmf $p(x)$ and expected value μ. Then the variance of X, denoted by $V(X)$ or σ_X^2, or just σ^2, is

$$V(X) = \sum_D (x-\mu)^2 \cdot p(x) = \mathbb{E}[(X-\mu)^2]$$

The number of arithmetic operations necessary to compute σ^2 can be reduced by using an alternative formula

$$V(X) = \sigma^2 = \left[\sum_D x^2 \cdot p(x)\right] - \mu_x^2 = E(X^2) - [E(X)]^2.$$

a.) The mean of the above pmf is given by

$$\begin{aligned}\mu_x &= \sum_{x \in D} xp(x) \\ &= 0(0.5) + 1(0.2) + 2(0.15) + 3(0.1) + 4(0.05) \\ &= 1.\end{aligned}$$

b.) The variance is given by

$$\begin{aligned}\mathrm{Var}(x) = E(X^2) - [E(X)]^2 &= \left[\sum_{x \in D} x^2 p(x)\right] - \mu_x^2 \\ &= \left[0^2(0.5) + 1^2(0.2) + 2^2(0.15) + 3^2(0.1) + 4^2(0.05)\right] - 1^2 \\ &= 2.5 - 1 = 1.5.\end{aligned}$$

Remark. $E[h(X)]$, where $h(X)$ is a linear function, is easily computed from $E(X)$.

$$E(h(X) = E(aX+b) = a \cdot E(X) + b$$

c.) Thus, we have

$$\begin{aligned}
&50E(X) - 20 \\
&= 50(1) - 20 \\
&= 30.
\end{aligned}$$

2. Suppose that X is a discrete random variable having the following probability mass function.

x	0	1	2	3	4	5
$p(x)$	0.05	0.20	0.40	0.15	0.10	0.10

(a) Find the mean value of X.

(b) Find the standard deviation σ.

(c) Suppose that $Y = 2X + 1$. Find each of the following.

(i) $E(Y)$

(ii) $\text{Var}(Y)$

(d) Suppose that $W = X^2 - 2X + 3$. Find $E(W)$.

a.) The expected value $E(X)$ is given by

$$\begin{aligned} E(X) &= \sum_{x \in D} xp(X) \\ &= 0(0.05) + 1(0.2) + 2(0.4) + 3(0.15) + 4(.10) + 5(.10) \\ &= 2.35. \end{aligned}$$

b.) The standard deviation σ is given by

$$\begin{aligned} \sigma x &= \sqrt{\text{Var}(X)} = \sqrt{E(X^2) - [E(X)]^2} \\ &= \sqrt{\sum_{x \in D} x^2 p(x) - \mu_x^2} \\ &= \sqrt{(0^2(0.05) + 1^2(0.2) + 2^2(0.4) + 3^2(0.15) + 4^2(0.1) + 5^2(0.1)) - 2.35^2} \\ &= \sqrt{7.25 - 5.5225} = \sqrt{1.7275} = 1.3143. \end{aligned}$$

c.i) The expected value $E(Y)$ is given by

$$\begin{aligned} E(Y) &= 2E(X) + 1 \\ &= 2(2.35) + 1 = 5.7. \end{aligned}$$

c.ii) For a linear transformation $Y = aX + B$, the variance is given by $a^2\text{Var}(X)$. Thus,

$$\begin{aligned} \text{Var}(Y) &= 2^2 \left(E(X^2) - [E(X)]^2\right) \\ &= 4\,(7.25 - 5.5225) \\ &= 6.91. \end{aligned}$$

d.) We need to find $E(W) = E(X^2 - 2X + 3)$, so we replace X by $X^2 - 2X + 3$ in $E(X) = \sum_{x \in D} xp(x)$ and derive some result

$$\begin{aligned} E(W) &= \sum_{x \in D} (x^2 - 2x + 3)p(x) \\ &= \sum_{x \in D} x^2 p(x) - 2xp(x) + 3p(x) \\ &= \sum_{x \in D} x^2 p(x) - 2\sum_{x \in D} xp(x) + 3\sum_{x \in D} p(x) \\ \therefore E(W) &= E(X^2) - 2E(X) + 3 = 7.25 - 2(2.35) + 3 = 5.55. \end{aligned}$$

3. Calculate each of the following Binomial probabilities directly from the probability mass function.

(a) $P(X = 2)$ where $X \sim \text{Bin}(n = 8, p = 0.40)$

(b) $P(1 < X < 4)$ where $X \sim \text{Bin}(8, 0.40)$

(c) $P(X \leqslant 1)$ where $X \sim \text{Bin}(4, 0.50)$

(d) $P(X = 6)$ where $X \sim \text{Bin}(7, 0.80)$

Remark. The pmf for a binomial experiment that satisfies

1. The experiment consists of a sequence of n smaller experiments called **trials** , where n is fixed in advance of the experiment.
2. Each trial can result in one of the same two possible outcomes (dichotomous trials), which we generically denote by success (S) and failure (F).
3. The trials are independent, so that the outcome on any particular trial does not influence the outcome on any other trial.
4. The probability of success $P(S)$ is constant from trial to trial; we denote this probability by p

Is given by

$$b(x; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{for } x = 0, 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

Whereas the cdf is given by

$$P(X = x) = B(x; n, p) = P(X \leqslant x) = \sum_{y=0}^{x} b(y; n, p) \ .$$

Probabilities $P(X \geqslant a)$ are given by $1 - B(a-; n, p)$, where $a-$ represents the largest x value strictly less than a

a.) By the binomial pmf, we have

$$\begin{aligned} b(2; 8, 0.4) &= \binom{n}{x} p^x (1-p)^{n-x} = \binom{8}{2} 0.4^2 \cdot 0.6^6 \\ &= \frac{8 \cdot 7 \cdot 6!}{2!6!} \cdot 0.4^2 \cdot 0.6^6 = \frac{56}{2} \cdot 0.4^2 \cdot 0.6^6 \\ &= 0.2090. \end{aligned}$$

b.) This probablity is computed by $P(X = 2) + P(X = 3) = b(2; 8, 0.4) + b(3; 8, 0.4)$. Thus,

$$\begin{aligned} P(1 < X < 4) &= \binom{8}{2} \cdot 0.4^2 \cdot 0.6^6 + \binom{8}{3} \cdot 0.4^3 \cdot 0.6^5 \\ &= 0.2090 + 0.2787 = 0.4877. \end{aligned}$$

c.) By the cdf, we have

$$
\begin{aligned}
P(X \leqslant 1) = B(1;4,0.5) &= \sum_{y=0}^{1} b(y;4,0.5) \\
&= b(0;4,0.5) + b(1;4,0.5) \\
&= \binom{4}{0} \cdot 0.5^0 \cdot 0.5^4 + \binom{4}{1} \cdot 0.5^1 \cdot 0.5^3 \\
&= 0.0625 + 0.25 = 0.3125.
\end{aligned}
$$

d.) By the pmf, we have

$$
\begin{aligned}
P(X = 6) = b(6;7,0.8) &= \binom{7}{6} \cdot 0.8^6 \cdot 0.2^1 \\
&= 0.367.
\end{aligned}
$$

4. A particular telephone number is used to receive both voice calls and fax messages. Suppose that 25% of the incoming calls involve fax messages. Let X = the number of fax messages among a random sample of 15 calls so that $X \sim \text{Bin}(15, 0.25)$.

(a) Use the table of cumulative Binomial probabilities to calculate the chance of each of the following events.

  (i) At most 3 of the calls involve fax messages
  (ii) Between 4 and 8 of the calls (inclusive of the endpoints) involve fax messages
  (iii) Exactly 4 of the calls involve fax messages
  (iv) $P(2 \leqslant X \leqslant 7)$

(b) Calculate the (i) mean and the (ii) standard deviation of the number of calls, out of 15, that would involve fax messages.

a.i)

$$P(X \leqslant 3) = 0.461.$$

a.ii)

$$\begin{aligned} P(4 \leqslant X \leqslant 8) &= B(8; 15, 0.25) - B(3; 15, 0.25) \\ &= 0.996 - 0.461 = 0.535. \end{aligned}$$

a.iii)

$$\begin{aligned} P(X = 4) &= B(4; 15, 0.25) - B(3; 15, 0.25) \\ &= 0.686 - 0.461 = 0.225. \end{aligned}$$

a.iv)

$$\begin{aligned} P(2 \leqslant X \leqslant 7) &= B(7; 15, 0.25) - B(1; 15, 0.25) \\ &= 0.983 - 0.080 = 0.903. \end{aligned}$$

Remark. If $X \sim \text{Bin}(n, p)$, then $E(X) = np$, $V(X) = np(1-p) = npq$, and $\sigma_X = \sqrt{npq}$ (where $q = 1 - p$).

b.)

$$\begin{aligned} \mu_x &= np = 15(0.25) = 3.75 \\ \sigma_x &= \sqrt{np(1-p)} = \sqrt{3.75(0.75)} = 1.6771. \end{aligned}$$

5. An individual who has automobile insurance from a certain company is randomly selected. Let $X =$ the number of moving violations for which the individual was cited during the last 3 years. The probability mass function of X is given below.

x	0	1	2	3	4
$p(x)$	0.50	0.20	0.15	0.10	0.05

(a) Suppose that we define an individual who has automobile insurance from the company as a 'success' if they have no moving violations. Calculate the probability that, in a random sample of 10 individuals, exactly 4 will have no moving violations.

(b) Suppose that we define an individual as a 'success' if they have at least two moving violations. Calculate the probability that at most 3 out of 10 people will each have at least two moving violations.

(c) Calculate the (i) mean and the (ii) standard deviation of the number of individuals out of 10 who have exactly one moving violation.

a.) We define $X \sim Bin(10, 0.5)$. This implies

$$\begin{aligned} b(4; 10, 0.5) &= \binom{10}{4} \cdot 0.5^4 \cdot 0.5^6 \\ &= 0.2051. \end{aligned}$$

b.) We define $X \sim Bin(10, 0.3)$. Thus,

$$B(3; 10, 0.3) = \sum_{y=0}^{3} b(y; 10, 0.3)$$

.

By table A.1, $B(3; 10, 0.3) = 0.650$

c.) Define $X \sim Bin(10, 0.2)$, the mean and standard deviation are

$$\begin{aligned} \mu_x &= np = 10(0.2) = 2 \\ \sigma_x &= \sqrt{np(1-p)} = \sqrt{2(0.8)} = 1.2649. \end{aligned}$$