

Selected Problems Chapter 3
Introduction to Probability for Data Science
Stanley Chan

Mustaf Ahmed

September 23, 2022

Problem 3.2. Two dice are tossed. Let X be the absolute difference in the number of dots facing up.

- a. Find and plot the PMF of X .
- b. Find the probability $X \leq 2$.
- c. Find $\mathbb{E}[X]$ and $Var[X]$.

Part a. The possible random variable states are $X(\Omega) = \{0, 1, 2, 3, 4, 5\}$. The probability mass function for this random variable is

$$p_X(0) = P(\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}) = \frac{1}{6}$$

$$p_X(1) = P(\{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3), (4, 5), (5, 4), (5, 6), (6, 5)\}) = \frac{5}{18}$$

$$p_X(2) = P(\{(1, 3), (3, 1), (2, 4), (4, 2), (3, 5), (5, 3), (4, 6), (6, 4)\}) = \frac{2}{9}$$

$$p_X(3) = P(\{(1, 4), (4, 1), (2, 5), (5, 2), (3, 6), (6, 3)\}) = \frac{1}{6}$$

$$p_X(4) = P(\{(1, 5), (5, 1), (2, 6), (6, 2)\}) = \frac{1}{9}$$

$$p_X(5) = P(\{(1, 6), (6, 1)\}) = \frac{1}{18}$$

Part b.

$$\begin{aligned}P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\&= \frac{2}{3}\end{aligned}$$

Part c.

$$\begin{aligned}\mathbb{E}[X] &= 0 * \frac{1}{6} + 1 * \frac{5}{8} + 2 * \frac{2}{9} + 3 * \frac{1}{6} + 4 * \frac{1}{9} + 5 * \frac{1}{18} \\&= \frac{35}{18}\end{aligned}$$

$$\begin{aligned}\text{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\&= \mathbb{E}[X^2] - \left(\frac{35}{18}\right)^2 \\&= 0^2 * \frac{1}{6} + 1^2 * \frac{5}{8} + 2^2 * \frac{2}{9} + 3^2 * \frac{1}{6} + 4^2 * \frac{1}{9} + 5^2 * \frac{1}{18} - \frac{35^2}{18^2} \\&= \frac{665}{324}\end{aligned}$$

Problem Theorem 3.4. Prove that the expectation of a random variable X has the following properties:

(a) **Function.** For any g ,

$$\mathbb{E}[g(X)] = \sum_{x \in X(\Omega)} g(x)p_X(x).$$

(b) **Linearity.** For any function g and h ,

$$\mathbb{E}[g(X) + h(X)] = \mathbb{E}[g(X)] + \mathbb{E}[h(X)].$$

(c) **Scale.** For any constant c ,

$$\mathbb{E}[cX] = c\mathbb{E}[X].$$

(d) **DC shift.** For any constant c ,

$$\mathbb{E}[X + c] = \mathbb{E}[X] + c$$

Part a. Let g be a function.

By the definition of expectation, we have

$$\mathbb{E}[g(X)] = \sum_{s \in g(X(\Omega))} sp_{g(X)}(s).$$

Expanding the the term $p_{g(X)}(s)$, we get

$$\begin{aligned} p_{g(X)}(s) &= P(\{e \in \Omega \mid g(X(e)) = s\}) \\ &= \sum_{x \in g^{-1}(s)} p_X(x). \end{aligned}$$

We can now simplify the original equation:

$$\begin{aligned} \mathbb{E}[g(X)] &= \sum_{s \in g(X(\Omega))} sp_{g(X)}(s) \\ &= \sum_{s \in g(X(\Omega))} s \sum_{x \in g^{-1}(s)} p_X(x) \\ &= \sum_{s \in g(X(\Omega))} \sum_{x \in g^{-1}(s)} sp_X(x) \\ &= \sum_{s \in g(X(\Omega))} \sum_{x \in g^{-1}(s)} g(x)p_X(x) \\ &= \sum_{x \in X(\Omega)} g(x)p_X(x), \end{aligned}$$

where the last relation comes summing over each part of the partition of $X(\Omega)$.

Part b. Let g, h be functions. By 3.4(a), we have

$$\begin{aligned}
 \mathbb{E}[g(X) + h(X)] &= \sum_{x \in X(\Omega)} (g(x) + h(x))p_X(x) \\
 &= \sum_{x \in X(\Omega)} g(x)p_X(x) + \sum_{x \in X(\Omega)} h(x)p_X(x) \\
 &= \mathbb{E}[g(X)] + \mathbb{E}[h(X)].
 \end{aligned}$$

Part c. Let g be a function, and let c be a constant. Define $h(X) = cX$. By 3.4(a), we have

$$\begin{aligned}
 \mathbb{E}[h(X)] &= \mathbb{E}[cX] \\
 &= \sum_{x \in X(\Omega)} h(x)p_X(x) \\
 &= \sum_{x \in X(\Omega)} cxp_X(x) \\
 &= c \sum_{x \in X(\Omega)} xp_X(x) \\
 &= c\mathbb{E}[X].
 \end{aligned}$$

Part d. Let g be a function, and let c be a constant. Define $h(X) = X + c$. By 3.4(a), we have

$$\begin{aligned}
 \mathbb{E}[h(X)] &= \mathbb{E}[X + c] \\
 &= \sum_{x \in X(\Omega)} h(x)p_X(x) \\
 &= \sum_{x \in X(\Omega)} (x + c)p_X(x) \\
 &= \sum_{x \in X(\Omega)} xp_X(x) + \sum_{x \in X(\Omega)} cp_X(x) \\
 &= \sum_{x \in X(\Omega)} xp_X(x) + c \sum_{x \in X(\Omega)} p_X(x) \\
 &= \sum_{x \in X(\Omega)} xp_X(x) + c \\
 &= \mathbb{E}[X] + c.
 \end{aligned}$$

Problem Theorem 3.5. The variance of a random variable X has the following properties:

(i) **Moment.**

$$\text{Var}[X] = E[X^2] - E[X]^2$$

(ii) **Scale.** For any constant c ,

$$\text{Var}[cX] = c^2 \text{Var}[X]$$

(iii) **DC Shift.** For any constant c ,

$$\text{Var}[X + c] = \text{Var}[X]$$

Proof. **Part (i).**

$$\begin{aligned} \text{Var}[X] &= E[(X - \mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - 2E[X]^2 + E[X]^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

Part (ii).

$$\begin{aligned} \text{Var}[cX] &= E[(cX - c\mu)^2] \\ &= E[c^2(X - \mu)^2] \\ &= c^2 E[(X - \mu)^2] \\ &= c^2 \text{Var}[X] \end{aligned}$$

Part (iii).

$$\begin{aligned} \text{Var}[X + c] &= E[(X + c - c - E[X])^2] \\ &= E[(X - E[X])^2] \\ &= \text{Var}[X] \end{aligned}$$

□