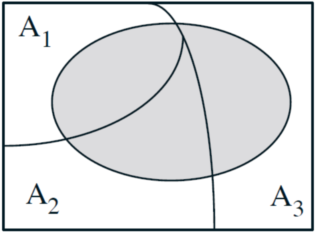
Total Expectation Theorem, Conditional Variance, Expectation of Geometric Distribution, Joint PMFs, Conditional Joint PMFs

Total Expectation Theorem

- Partition of sample space into disjoint events $A_1, A_2, \cdots, A_n$



$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B \mid A_n)$$

$$p_X(x) = \mathbf{P}(A_1)p_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)p_{X|A_n}(x)$$

$$\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X \mid A_n]$$

Total Expectation Theorem—Example

- Given are the following events:

- $A_1 = \{1, 4, 7, 10\}$
- $A_2 = \{2, 5, 8\}$
- $A_3 = \{3, 6, 9\}$

$$P(A_1) = \frac{1}{2}, P(A_2) = \frac{1}{3}, P(A_3) = \frac{1}{6}$$

- We choose a random number between 1 to 10 with the given probabilities. Let random variable X be the outcome. What is $E[X]$?

- Using total expectation theorem:

$$E[X] = P(A_1)E[X|A_1] + P(A_2)E[X|A_2] + P(A_3)E[X|A_3]$$

- We should compute $E[X|A_1], E[X|A_2]$ and $E[X|A_3]$.

Total Expectation Theorem—Example

$$E[X|A_1] = 1 \cdot p_{X|A_1}(1) + 4 \cdot p_{X|A_1}(4) + 7 \cdot p_{X|A_1}(7) + 10 \cdot p_{X|A_1}(10) =$$

$$= (1 + 4 + 7 + 10)\frac{1}{4} = 5.5$$

$$E[X|A_2] = 2 \cdot p_{X|A_2}(2) + 5 \cdot p_{X|A_2}(5) + 8 \cdot p_{X|A_2}(8) =$$

$$= (2 + 5 + 8)\frac{1}{3} = 5$$

$$E[X|A_3] = 3 \cdot p_{X|A_3}(3) + 6 \cdot p_{X|A_3}(6) + 9 \cdot p_{X|A_3}(9) =$$

$$= (3 + 6 + 9)\frac{1}{3} = 6$$

Therefore,

$$E[X] = \frac{1}{2} \cdot (5.5) + \frac{1}{3} \cdot 5 + \frac{1}{6} \cdot 6 = \frac{16.25}{3} \approx 5.42$$

Total Expectation Theorem—Example

- Alternative solution (if we don't use total expectation theorem):

$$p_X(x) = P(A_1)p_{X|A_1}(x) + P(A_2)p_{X|A_2}(x) + P(A_3)p_{X|A_3}(x)$$

  if $i \in \{1, 4, 7, 10\}$, then:

$$p_{X|A_1}(i) = \frac{1}{4}, p_{X|A_2}(i) = 0, p_{X|A_3}(i) = 0$$

$$\Rightarrow p_X(i) = P(A_1)p_{X|A_1}(i) = \frac{1}{2} * \frac{1}{4} = \frac{1}{8}$$

$$\text{So,} \quad p_X(1) = p_X(4) = p_X(7) = p_X(10) = \frac{1}{8}$$

- If $i \in \{2, 5, 8\}$, then:

$$p_{X|A_1}(i) = 0, p_{X|A_2}(i) = \frac{1}{3}, p_{X|A_3}(i) = 0$$

$$\Rightarrow p_X(i) = P(A_2)p_{X|A_2}(i) = \frac{1}{3} * \frac{1}{3} = \frac{1}{9}$$

$$\text{So,} \quad p_X(2) = p_X(5) = p_X(8) = \frac{1}{9}$$

Total Expectation Theorem—Example

- if $i \in \{3, 6, 9\}$, then:

$$p_{X|A_1}(i) = 0, p_{X|A_2}(i) = 0, p_{X|A_3}(i) = \frac{1}{3}$$

$$\Rightarrow p_X(i) = P(A_3)p_{X|A_3}(i) = \frac{1}{6} * \frac{1}{3} = \frac{1}{18}$$

$$\text{So,} \quad p_X(3) = p_X(6) = p_X(9) = \frac{1}{18}$$

$$E[X] = \sum_x x p_X(x) =$$

$$= (1 + 4 + 7 + 10)\frac{1}{8} + (2 + 5 + 8)\frac{1}{9} + (3 + 6 + 9)\frac{1}{18} = \frac{16.25}{3} \approx 5.42$$

Conditional Variance

$$Var[X|A] = E[X^2|A] - (E[X|A])^2$$

In the previous example, what is $Var[X|A_1]$?

- We know $E[X|A_1] = 5.5$
- Let's compute $E[X^2|A_1]$:

$$X = \{1, 4, 7, 10\} \Rightarrow X^2 = \{1, 16, 49, 100\}$$

$$P(X^2 = k^2|A_1) = p_{X^2|A_1}(k^2) = p_{X|A_1}(k) = P(X = K|A_1)$$

$$If \;\; k \in \{1, 4, 7, 10\} \;\; \text{then} \;\; p_{X|A_1}(k) = \frac{1}{4} \Rightarrow p_{X^2|A_1}(k^2) = \frac{1}{4}$$

$$E[X^2|A_1] = \frac{(1 + 16 + 49 + 100)}{4} = 41.5$$

- Therefore,

$$Var[X|A_1] = E[X^2|A_1] - (E[X|A_1])^2 = 41.5 - (5.5)^2 = 11.25$$

Geometric Distribution—Expectation

You have a biased coin, $P(Head) = p$ and $P(Tail) = 1 - p$, and $p > 0$. You toss the coin until you see a "Head" and then you stop. Assume tosses are independent. Define random variable X to be "number of coin tosses until first head". Then

$$p_X(k) = P(X = k) = P(\text{“number of coin tosses until first head''} = k)$$

Event A_1 and A_2 partition the sample space and they are disjoint.

$$A_1 = \{X = 1\} \text{ and } A_2 = \{X > 1\}$$

- Using total expectation theorem:

$$E[X] = P(A_1)E[X|A_1] + P(A_2)E[X|A_2]$$

- $P(A_1) = P(X = 1) = P(\text{head in the first toss}) = p \Rightarrow p(A_2) = 1 - p$
- $E[X|A_1] = E[X|X = 1] = 1$

Geometric Distribution—Expectation

- $E[X|A_2] = E[X|X > 1]$
- Linearity of expectation: $E[X|X > 1] = E[X - 1|X > 1] + E[1|X > 1]$
- More generally, using linearity of expectation, for event A and a constant c:

$$E[X|A] = E[X - c|A] + E[c|A]$$

  and $E[c|A] = c$, therefore:

$$\boxed{E[X|A] = E[X - c|A] + c}$$

- So $E[X|A_2] = E[X|X > 1] = E[X - 1|X > 1] + 1$
- Using **memoryless property** of geometric distribution:

$$E[X] = E[X - 1|X > 1]$$

- So $E[X|A_2] = E[X] + 1$.
- $E[X] = P(A_1)E[X|A_1] + P(A_2)E[X|A_2]$. Therefore,

$$E[X] = p \cdot 1 + (1 - p)(E[X] + 1)$$

$$\Rightarrow \boxed{E[X] = \frac{1}{p}}$$

Geometric Distribution—Expectation

- For Geometric distribution with parameter p:

$$\boxed{E[X] = \frac{1}{p}}$$

- You have a fair coin $p = \frac{1}{2}$. What is the expected number of tosses to see a head?

$$E[X] = \frac{1}{p} = 2$$

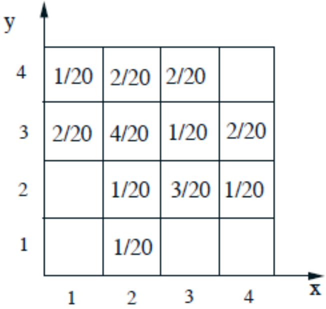
- You have a coin and $P(\text{head}) = \frac{1}{3}$. What is the expected number of tosses to see a head?

$$E[X] = \frac{1}{p} = 3$$

Joint PMFs

- Definition:

$$p_{X,Y}(x,y) = P(X = x \text{ and } Y = y)$$



- $p_{X,Y}(2,3) = \frac{4}{20}$, $p_{X,Y}(1,1) = 0$
- Covering the entire sample space (here the square):

$$\sum_x \sum_y p_{X,Y}(x,y) = 1$$

Joint PMFs

Obtaining PMF of X or Y from joint PMF of X and Y:

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

y \ x	1	2	3	4
4	1/20	2/20	2/20	
3	2/20	4/20	1/20	2/20
2		1/20	3/20	1/20
1		1/20		

$$p_X(3) = \sum_y p_{X,Y}(3, y) = p_{X,Y}(3,1) + p_{X,Y}(3,2) + p_{X,Y}(3,3) + p_{X,Y}(3,4) =$$

$$= 0 + \frac{3}{20} + \frac{1}{20} + \frac{2}{20} = \frac{6}{20}$$

- $p_Y(4) = \sum_x p_{X,Y}(x, 4) = p_{X,Y}(1,4) + p_{X,Y}(2,4) + p_{X,Y}(3,4) + p_{X,Y}(4,4) =$
  $$= \frac{1}{20} + \frac{2}{20} + \frac{2}{20} + 0 = \frac{5}{20}$$

Joint PMFs—Conditional Probability

- Definition:

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{P(X = x \text{ and } Y = y)}{p_Y(y)}$$

- $p_{X|Y}(x|2) = ?$

$$p_{X|Y}(1 \mid 2) = \frac{p_{X,Y}(1,2)}{p_Y(2)} = \frac{0}{5/20} = 0$$

$$p_{X|Y}(2 \mid 2) = \frac{p_{X,Y}(2,2)}{p_Y(2)} = \frac{1/20}{5/20} = \frac{1}{5}$$

$$p_{X|Y}(3 \mid 2) = \frac{p_{X,Y}(3,2)}{p_Y(2)} = \frac{3/20}{5/20} = \frac{3}{5}$$

$$p_{X|Y}(4 \mid 2) = \frac{p_{X,Y}(4,2)}{p_Y(2)} = \frac{1/20}{5/20} = \frac{1}{5}$$

- Also, in general: $\sum_x p_{X|Y}(x|y) = 1$
- $\sum_x p_{X|Y}(x|2) = p_{X|Y}(1|2) + p_{X|Y}(2|2) + p_{X|Y}(3|2) + p_{X|Y}(4|2) = 1$

Example

Random variables X and Y have the joint PMF

$$p_{X,Y}(x,y) = \begin{cases} c(x^2+y^2), & \text{if } x \in \{1,2,4\} \text{ and } y \in \{1,3\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) What is the value of the constant c?

(b) What is $\mathbf{P}(Y < X)$?

(c) What is $\mathbf{P}(Y > X)$?

(d) What is $\mathbf{P}(Y = X)$?

(e) What is $\mathbf{P}(Y = 3)$?

(f) Find the marginal PMFs $p_X(x)$ and $p_Y(y)$.

(g) Find the expectations $\mathbf{E}[X]$, $\mathbf{E}[Y]$ and $\mathbf{E}[XY]$.

(h) Find the variances $\text{var}(X)$, $\text{var}(Y)$ and $\text{var}(X+Y)$.

(i) Let A denote the event $X \geq Y$. Find $\mathbf{E}[X \mid A]$ and $\text{var}(X \mid A)$.

Example

(a) From the joint PMF, there are six $(x,\ y)$ coordinate pairs with nonzero probabilities of occurring. These pairs are (1, 1), (1, 3), (2, 1), (2, 3), (4, 1), and (4, 3). The probability of a pair is proportional to the sum of the squares of the coordinates of the pair, $x^2 + y^2$. Because the probability of the entire sample space must equal 1, we have:

$$(1+1)c + (1+9)c + (4+1)c + (4+9)c + (16+1)c + (16+9)c = 1.$$

Solving for c, we get $c = \boxed{\frac{1}{72}}$.

(b) There are three sample points for which $y < x$:

$$\mathbf{P}(Y < X) = \mathbf{P}\left(\{(2,1)\}\right) + \mathbf{P}\left(\{(4,1)\}\right) + \mathbf{P}\left(\{(4,3)\}\right) = \frac{5}{72} + \frac{17}{72} + \frac{25}{72} = \boxed{\frac{47}{72}}.$$

(c) There are two sample points for which $y > x$:

$$\mathbf{P}(Y > X) = \mathbf{P}\left(\{(1,3)\}\right) + \mathbf{P}\left(\{(2,3)\}\right) = \frac{10}{72} + \frac{13}{72} = \boxed{\frac{23}{72}}.$$

Example

(d) There is only one sample point for which $y = x$:

$$\mathbf{P}(Y = X) = \mathbf{P}(\{(1,1)\}) = \boxed{\tfrac{2}{72}}.$$

Notice that, using the above two parts,

$$\mathbf{P}(Y < X) + \mathbf{P}(Y > X) + \mathbf{P}(Y = X) = \frac{47}{72} + \frac{23}{72} + \frac{2}{72} = 1$$

as expected.

(e) There are three sample points for which $y = 3$:

$$\mathbf{P}(Y = 3) = \mathbf{P}(\{(1,3)\}) + \mathbf{P}(\{(2,3)\}) + \mathbf{P}(\{(4,3)\}) = \frac{10}{72} + \frac{13}{72} + \frac{25}{72} = \boxed{\tfrac{48}{72}}.$$

Example

(f) In general, for two discrete random variable X and Y for which a joint PMF is defined, we have

$$p_X(x) = \sum_{y=-\infty}^{\infty} p_{X,Y}(x,y) \qquad \text{and} \qquad p_Y(y) = \sum_{x=-\infty}^{\infty} p_{X,Y}(x,y).$$

In this problem the ranges of X and Y are quite restricted so we can determine the marginal PMFs by enumeration. For example,

$$p_X(2) = \mathbf{P}\left(\{(2,1)\}\right) + \mathbf{P}\left(\{(2,3)\}\right) = \frac{18}{72}\,.$$

Overall, we get:

$$p_X(x) = \begin{cases} 12/72, & \text{if } x = 1, \\ 18/72, & \text{if } x = 2, \\ 42/72, & \text{if } x = 4, \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad p_Y(y) = \begin{cases} 24/72, & \text{if } y = 1, \\ 48/72, & \text{if } y = 3, \\ 0, & \text{otherwise.} \end{cases}$$

Example

(g) In general, the expected value of any discrete random variable X equals

$$\mathbf{E}[X] = \sum_{x=-\infty}^{\infty} x p_X(x).$$

For this problem,

$$\mathbf{E}[X] = 1 \cdot \frac{12}{72} + 2 \cdot \frac{18}{72} + 4 \cdot \frac{42}{72} = \boxed{3}$$

and

$$\mathbf{E}[Y] = 1 \cdot \frac{24}{72} + 3 \cdot \frac{48}{72} = \boxed{\tfrac{7}{3}}.$$

To compute $\mathbf{E}[XY]$, note that $p_{X,Y}(x,y) \neq p_X(x)p_Y(y)$. Therefore, X and Y are not independent and we cannot assume $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$. Thus, we have

$$\begin{aligned}
\mathbf{E}[XY] &= \sum_x \sum_y xy p_{X,Y}(x,y) \\
&= 1 \cdot \frac{2}{72} + 2 \cdot \frac{5}{72} + 4 \cdot \frac{17}{72} + 3 \cdot \frac{10}{72} + 6 \cdot \frac{13}{72} + 12 \cdot \frac{25}{72} = \boxed{\tfrac{61}{9}}.
\end{aligned}$$

Example

(h) The variance of a random variable X can be computed as $\mathbf{E}[X^2]-\mathbf{E}[X]^2$ or as $\mathbf{E}[(X-\mathbf{E}[X])^2]$. We use the second approach here because X and Y take on such limited ranges. We have

$$\mathrm{var}(X) = (1-3)^2\frac{12}{72} + (2-3)^2\frac{18}{72} + (4-3)^2\frac{42}{72} = \boxed{\tfrac{3}{2}}$$

and

$$\mathrm{var}(Y) = (1-\frac{7}{3})^2\frac{24}{72} + (3-\frac{7}{3})^2\frac{48}{72} = \boxed{\tfrac{8}{9}}\,.$$

X and Y are not independent, so we cannot assume $\mathrm{var}(X+Y) = \mathrm{var}(X)+\mathrm{var}(Y)$. The variance of $X+Y$ will be computed using $\mathrm{var}(X+Y) = \mathbf{E}[(X+Y)^2]-(\mathbf{E}[X+Y])^2$. Therefore, we have

$$\mathbf{E}[(X+Y)^2] = 4\cdot\frac{2}{72} + 9\cdot\frac{5}{72} + 25\cdot\frac{17}{72} + 16\cdot\frac{10}{72} + 25\cdot\frac{13}{72} + 49\cdot\frac{25}{72} = \frac{547}{18}\,.$$

$$(\mathbf{E}[X+Y])^2 = (\mathbf{E}[X]+\mathbf{E}[Y])^2 = \left(3+\frac{7}{3}\right)^2 = \frac{256}{9}\,.$$

Therefore,

$$\mathrm{var}(X+Y) = \frac{547}{18} - \frac{256}{9} = \boxed{\tfrac{35}{18}}\,.$$

Example

(i) There are four $(x,\ y)$ coordinate pairs in A : (1,1), (2,1), (4,1), and (4,3). Therefore, $\mathbf{P}(A) = \frac{1}{72}(2+5+17+25) = \frac{49}{72}$. To find $\mathbf{E}[X \mid A]$ and $\text{var}(X \mid A)$, $p_{X|A}(x)$ must be calculated. We have

$$p_{X|A}(x) = \begin{cases} 2/49, & \text{if } x = 1, \\ 5/49, & \text{if } x = 2, \\ 42/49, & \text{if } x = 4, \\ 0, & \text{otherwise}, \end{cases}$$

$$\begin{aligned}
\mathbf{E}[X \mid A] &= 1 \cdot \frac{2}{49} + 2 \cdot \frac{5}{49} + 4 \cdot \frac{42}{49} = \boxed{\frac{180}{49}}, \\
\mathbf{E}[X^2 \mid A] &= 1^2 \cdot \frac{2}{49} + 2^2 \cdot \frac{5}{49} + 4^2 \cdot \frac{42}{49} = \frac{694}{49}, \\
\text{var}(X \mid A) &= \mathbf{E}[X^2 \mid A] - (\mathbf{E}[X \mid A])^2 = \frac{694}{49} - \left(\frac{180}{49}\right)^2 = \boxed{\frac{1606}{2401}},
\end{aligned}$$