Random Variables

Example

Each time a component is tested, the trial is a success (S) or a failure (F). Suppose the component is tested repeatedly until a success occurs on three consecutive trials. Let Y denote the number of trials necessary to achieve this. What possible values can Y take? List all outcomes ω mapping to the five smallest possible values of Y.

The possible values of Y are: $Y \in \{3, 4, 5, \ldots\}$.

$Y = 3$

- $\omega = SSS$

$Y = 4$

- $\omega = FSSS$

$Y = 5$

- $\omega = FFSSS$
- $\omega = SFSSS$

$Y = 6$

- $\omega = FFFSSS$
- $\omega = FSFSSS$
- $\omega = SFFSSS$
- $\omega = SSFSSS$

$Y = 7$

- $\omega = FFFFSSS$
- $\omega = FFSFSSS$
- $\omega = FSFFSSS$
- $\omega = FSSFSSS$
- $\omega = SFFFSSS$
- $\omega = SFSFSSS$
- $\omega = SFFFSSS$

Example

For each random variable X, describe the set of possible values for the variable. Is the variable discrete or continuous?

(a) The number of unbroken eggs in a randomly chosen one dozen carton

(b) The number of students absent in a particular class on a certain day

(c) The number of times a batter at bat swings to hit a ball at baseball

(d) The length of a randomly chosen rattlesnake

(e) The amount of earned from selling a 10-print set of woodcut

(f) The pH of a randomly chosen soil sample

(g) The tension (psi) of a randomly selected tuned guitar string

(h) The total number of coin tosses for three individuals to all match

---

(a) Discrete

(b) Discrete

(c) Discrete

(d) Continuous

(e) Continuous

(f) Continuous

(g) Continuous

(h) Discrete

Definition : Probability Distribution for Discrete Random Variables

Suppose that $X : \Omega \to \mathbb{R}$ is a discrete random variable. The probability distribution, or probability mass function (pmf) of X is a function p defined for every real number x by

$$p(x) = \Pr(X = x) = \Pr(\{\omega \in \Omega \mid X(\omega) = x\})$$

That is, a pmf specifies the probability of observing a specific value x for the rv X. Any pmf satisfies $p(x) \geq 0$ and $\sum_{x \in \mathbb{R}} p(x) = 1$.

Example

Five tickets, two red and three white, are in a hat. The tickets will be drawn at random until a red ticket is drawn. Let X be the number of draws required to draw a red ticket. What is the pmf p of X?

$$X \in \{1, 2, 3, 4\}$$

We define $p(x)$ for the possible values:

$$\begin{aligned}
p(1) &= \Pr(X = 1) = \frac{2}{5} = 0.4 \\
p(2) &= \Pr(X = 2) = \frac{3}{5} \cdot \frac{2}{4} = 0.3 \\
p(3) &= \Pr(X = 3) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = 0.2 \\
p(4) &= \Pr(X = 4) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{2}{2} = 0.1
\end{aligned}$$

$$p(x) = \begin{cases} 0.4 & \text{if} \quad x = 1 \\ 0.3 & \text{if} \quad x = 2 \\ 0.2 & \text{if} \quad x = 3 \\ 0.1 & \text{if} \quad x = 4 \end{cases}$$

We can also find the pmf of an infinite discrete rv X.

Example

Suppose that a 70% heads-loaded unfair coin is tossed until a head appears.

$$\Omega = \{H, TH, TTH, TTTH, \dots\}$$

Let X be the number of tosses observed. What is the pmf of X? Using the pmf, what are $\Pr(X = 7)$, $\Pr(X \leq 2)$ and $\Pr(X \leq 2.4)$?

$\Pr(\text{heads}) = 0.7$, and $\Pr(\text{tails}) = 0.3$. $X = k$ means the first $k - 1$ tosses are tails, and the kth toss is heads.

$$\Pr(X = k) = \Pr(TT \dots TH)$$

Also, all of the tosses are independent.

$$p(x) = \Pr(X = k) = \Pr(T)^{k-1} \Pr(H) = 0.3^{k-1} \cdot 0.7 \quad \text{for } k = 1, 2, \dots$$

So then,

$$Pr(X = 7) = 0.3^6 \cdot 0.7 = 0.0005103$$

$$\begin{aligned} \Pr(X \leq 2) &= \Pr(X = 1) + \Pr(X = 2) \\ &= 0.3^0 \cdot 0.7 + 0.3^1 \cdot 0.7 \\ &= 0.7 + 0.21 \\ &= 0.91 \end{aligned}$$

$$\Pr(X \leq 2.4) = \Pr(X \leq 2) = 0.91$$

Definition : Cumulative Distribution Function

The cumulative distribution function (cdf), F of a discrete rv X with pmf $p$is the function defined on all real numbers by

$$F(x) = \Pr(X \leq x) = \sum_{y \leq x} p(y)$$

For any x, $F(x)$ is the probability that X is at most x.

Example

Suppose that a 70% heads-loaded unfair coin is tossed until a head appears.

$$\Omega = \{H, TH, TTH, TTTH, \dots\}$$

Let X be the number of tosses observed. What is the cdf of X? Using the pmf, what are $\Pr(X \leq 10)$ and $\Pr(2 \leq X \leq 4)$?

$$\begin{aligned}
F(k) &= \Pr(X \leq k) \\
&= \sum_{i=1}^{k} \Pr(X = i) \\
&= \sum_{i=1}^{k} (1-P)^{i-1} \cdot P \\
&= P \cdot \sum_{i=1}^{k} (1-P)^{i-1} \qquad \text{Let } j = i - 1 \\
&= P \cdot \sum_{j=0}^{k-1} (1-P)^{j}
\end{aligned}$$

For $r = 1 - P$, we get: $F(k) = P\left(\frac{1-(1-P)^k}{1-(1-P)}\right) = 1 - (1-P)^k$. Hence,

$$F(k) = 1 - 0.3^k$$

$$\Pr(X \leq 10) = F(10) = 1 - 0.3^10 = 0.9999940951$$

$$\begin{aligned}
\Pr(2 \leq X \leq 4) &= F(4) - F(1) \\
&= \left(1 - 0.3^4\right) - (1 - 0.3) \\
&= 0.9919 - 0.7 \\
&= 0.2919
\end{aligned}$$

Knowing just the pmf gives enough information to compute probabilities about X. It is even possible without knowing the sample space Ω.

Example

Suppose that X is a random variable with pmf $p(x) = c\lambda^x/x!$ for $x = 0, 1, 2, \ldots$ where λ is some fixed positive value. What is the value of c? What are $\Pr(X = 0)$ and $\Pr(X > 2)$?

$$\sum_{x=0}^{\infty} \frac{c\lambda^x}{x!} = 1$$
$$c \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = 1$$

This infinite sum is exactly the Taylor series for e^λ.

$$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^\lambda$$

$$\begin{aligned} c \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} &= 1 \\ ce^\lambda &= 1 \\ c &= \frac{1}{e^\lambda} \\ c &= e^{-\lambda} \end{aligned}$$

$$p(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!} \qquad \text{This is a Poisson Distribution, } X \sim \text{Poisson}(\lambda)$$

$$\Pr(X = 0) = e^{-\lambda}$$

$$\begin{aligned} \Pr(X > 2) &= 1 - \Pr(X \leq 2) \\ &= 1 - (\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)) \\ &= 1 - \left(e^{-\lambda} + e^{-\lambda} \cdot \lambda + e^{-\lambda} \cdot \frac{\lambda^2}{2} \right) \\ &= 1 - e^{-\lambda} - e^{-\lambda} \cdot \lambda - e^{-\lambda} \cdot \frac{\lambda^2}{2} \end{aligned}$$