Theory of Probability Oct 21, 2020

Geometric random variable

↳ governs the likelihood of n failures, then one success.

Negative binomial random variable

↳ governs the number of trials needed to obtain k successes

Hypergeometric Random variable

For an urn with m white ball
n red balls

$N = m+n$ total balls.

Choose k balls at random, how many of the k balls are white.

Linearity of Expectation

One can show that for any two random variables X, Y,

$$E[X+Y] = E[X] + E[Y].$$

This is not true for variances:

$$\mathrm{Var}[\underbrace{X+Y}_{Z}] = \mathrm{Var}[Z]$$
$$= E[Z^2] - \left(E[Z]\right)^2$$

$$\Rightarrow \mathrm{Var}[X+Y] = E[(X+Y)^2] - (E[X+Y])^2$$

$$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$$

$$= E[X^2] + 2E[XY] + E[Y^2]$$

$$- E[\cancel{X}]^2 - 2E[X]E[Y] - E[Y]^2$$

$$= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2$$

$$+ 2(E[XY] - E[X]E[Y])$$

$$= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2(E[XY] - E[X]E[Y]).$$

$$\Rightarrow \mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

if and only if $E[XY] = E[X]E[Y]$.

Hypergeometric R.V.

$X \sim \text{Hypergeometric}(N, m, n)$

Urn with N balls, m are red, $N-m$ are blue.

Choose n balls at random.

Let X = # of red balls obtained.

$$P[X=k] = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$$

$X \sim$ roll of die 1

$Y \sim$ roll of die 2

$Z = X + Y$

$$E[Z] = \sum_{k=2}^{12} k\, P[Z=k]$$

$$= \sum_{k=2}^{12} k \sum_{i} P[X=i,\ Y=k-i]$$

Theoretical Example 4.37

Suppose X can be $x_1, x_2 \ldots \quad x_i$

Y can be $y_1, y_2, y_3 \ldots \quad y_j$

$X+Y=Z$ can take $z_1, z_2 \ldots \quad z_k$.

A_k = pairs of indices (i,j) such that $z_k = x_i + y_j$

$$= \{(i,j) : x_i + y_j = z_k\}$$

(a) Argue that

$$P[X+Y=z_k] = \sum_{(i,j)\in A_k} P[X=x_i,\ Y=y_j]$$

(b) Show that

$$E[\underbrace{X+Y}_{Z}] = \sum_{k} \sum_{(i,j)\in A_k} (x_i+y_j)\, P[X=x_i, Y=y_j]$$

$$E[Z] = \sum_k z_k P[Z = z_k].$$

$$= \sum_k z_k \sum_{(i,j) \in A_k} P[X = x_i, Y = y_j]$$

$$= \sum_k \sum_{(i,j) \in A_k} z_k P[X = x_i, Y = y_j]$$

$$= \sum_k \sum_{\substack{(i,j) \\ \in A_k}} (x_i + y_j) P[X = x_i, Y = y_j]. \quad \checkmark$$

(c) Using (b), show that

$$E[X+Y] = \sum_i \sum_j (x_i + y_j) P[X = x_i, Y = y_j]$$

Since each pair of indices (i,j) can only appear in exactly one A_k, we have that

$$\sum_k \sum_{\substack{(i,j) \\ \in A_k}} (x_i + y_j) P[X = x_i, Y = y_j]$$

$$= \sum_i \sum_j (x_i + y_j) P[X = x_i, Y = y_j]$$

$$= E[X+Y].$$

(d) Show that

$$P[X = x_i] = \sum_j P[X = x_i, Y = y_j].$$

$$E_{ij} = \{X = x_i, Y = y_j\}$$

$$\Rightarrow \{X = x_i\} = \bigcup_j E_{ij} \quad , \text{ also } E_{ij} \cap E_{ik} = \emptyset$$

unless $j = k$.

$$\Rightarrow P[X = x_i] = P[\bigcup_j E_{ij}]$$

$$= \sum P[E_{ij}]$$

$$= \sum_j P[X = x_i, Y = y_j].$$

$$\Rightarrow P[Y = y_j] = \sum_i P[X = x_i, Y = y_j].$$

(e) $$E[X+Y] = \sum_i \sum_j (x_i + y_j) P[X = x_i, Y = y_j]$$

$$= \sum_i \sum_j x_i P[X = x_i, Y = y_j]$$

$$+ \sum_i \sum_j y_j P[X = x_i, Y = y_j].$$

$$= \sum_i x_i \sum_j P[X = x_i, Y = y_j]$$

$$+ \sum_j y_j \sum_i P[X = x_i, Y = y_j]$$

$$= \underbrace{\sum_i x_i P[X = x_i]}_{E[X]} + \underbrace{\sum_j y_j P[Y = y_j]}_{E[Y]}.$$

$$= E[X] + E[Y].$$