

Recall

Independence

• A, B events $P(A \cap B) = P(A) P(B)$

"occurrence of A or B does not influence the probability of B or A respectively".

$n \geq 2$
 2^n equations } • A_1, \dots, A_n all mutually independent if

$$P\left(\bigwedge_{i=1}^n E_i\right) = \prod_{i=1}^n P(E_i) \quad \text{where } E_i = A_i^c \text{ or } A_i$$

• X, Y are two random variables (two)
they are independent if

$$P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$$

A, B are any two events

- discrete ; reduction : independent iff

$$P(X=x, Y=y) = P(X=x) P(Y=y)$$

$x \in \text{Range}(X)$
 $y \in \text{Range}(Y)$

Recall:

X, Y are two discrete r.v.

- Conditional distribution of $X|Y=y$

$$\{P(X=x|Y=y)\}_{x \in \text{Range}(X)}$$

- Conditional expectation of $X|Y=y$

$$E[X|Y=y] = \sum_{x \in \text{Range}(X)} x P(X=x|Y=y)$$

- Conditional variance of $X|Y=y$

$$E[(X - E[X|Y=y])^2 | Y=y] \equiv \text{Var}[X|Y=y]$$

Recall: Joint distribution - Discrete random variables

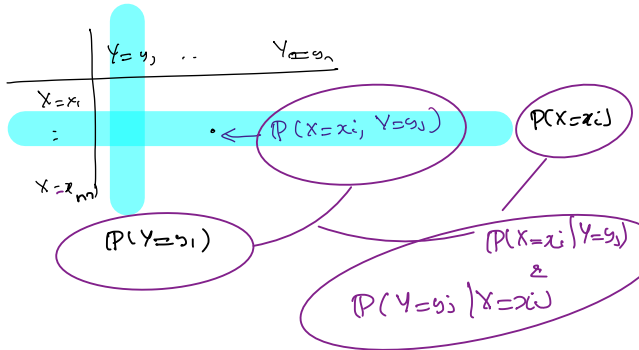
- X_1, X_2, \dots, X_n - Discrete random variables

Joint distribution of X_1, \dots, X_n is given by

$$\left\{ P(X_1 = x_1, \dots, X_n = x_n) \right\} \quad \begin{array}{l} x_i \in \text{Range}(X_i) \\ 1 \leq i \leq n \end{array}$$

- $n=2$

Row
sum



Multinomial

$k=2$ — corresponds to Binomial

I.I.D :- independent and identically distributed

Each trial or r.v.
has no effect on
others

They have the same
distribution

- Suppose we perform n i.i.d. trials each of which has k different possible outcomes.

n trials

x x x x ... x

} k outcomes

- For $j = 1, 2, \dots, k$, let p_j represent the probability any given trial results in the j^{th} outcome and let X_j represent the number of the n trials that result in the j^{th} outcome.

$$X_j = \sum_{m=1}^n \mathbb{1}_{\{\text{Trial } m \text{ has outcome } j\}}$$

- The joint distribution of all of the random variables X_1, X_2, \dots, X_k is called a “multinomial distribution”.

Multinomial - Joint distribution of (X_1, X_2, \dots, X_k)

- $0 \leq x_j \leq n \quad j = 1, 2, \dots, k \quad \text{and} \quad \sum_{j=1}^k x_j = n$
- Let $B(x_1, x_2, \dots, x_k) = \{X_1 = x_1, X_2 = x_2, \dots, X_k = x_k\}$.

Then,

$$P(B(x_1, x_2, \dots, x_k)) = \sum_{\omega \in B(x_1, x_2, \dots, x_k)} P(\{\omega\})$$

- Each $\omega \in B(x_1, x_2, \dots, x_k)$

$$P(\{\omega\}) = \prod_{j=1}^k (p_j)^{x_j}$$

and

$$|B(x_1, x_2, \dots, x_k)| = \frac{n!}{x_1! x_2! \dots x_k!}$$

Ex \equiv $\left\{ \begin{array}{l} \# \text{ of ways} \\ \text{of} \\ \text{allocating} \\ n \text{ balls} \\ \text{into} \\ k \text{ boxes.} \\ \text{Such that} \\ x_j \text{ of them are} \\ \text{in box } j \end{array} \right.$

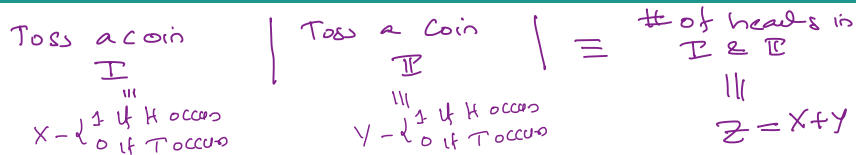
Multinomial

The joint distribution of X_1, X_2, \dots, X_k is given by

$$P(X_1 = x_1, \dots, X_k = x_k) = \begin{cases} \frac{n!}{x_1! x_2! \dots x_k!} \prod_{j=1}^k p_j^{x_j} & \text{if } x_j \in \{0, 1, \dots, n\} \\ & \text{and } \sum_{j=1}^k x_j = n \\ 0 & \text{otherwise} \end{cases}$$

Q:- How to work with independent random variables?

Sums of Random Variables - Independent



Let $X, Y \sim \text{Bernoulli}(p)$ be two independent random variables.

$Z = X + Y, Z \sim ?$.

$$\text{Range}(X) = \{0, 1\}$$

$$\text{Range}(Y) = \{0, 1\}$$

$$\text{Range}(Z) = \{0, 1, 2\}$$

$$\begin{aligned} \mathbb{P}(Z=0) &= \mathbb{P}(X+Y=0) = \mathbb{P}(X=0, Y=0) \\ &\stackrel{\text{independence}}{=} \mathbb{P}(X=0) \mathbb{P}(Y=0) \\ &= (1-p)(1-p) \\ &= (1-p)^2 \end{aligned}$$

$$\begin{aligned}
P(Z=1) &= P(X+Y=1) \\
&= P(\underbrace{X=0, Y=1 \cup X=1, Y=0}_{\text{mutually exclusive}}) \\
&= \underbrace{P(X=0, Y=1)} + \underbrace{P(X=1, Y=0)} \\
&\quad \text{independence} \\
&= P(X=0)P(Y=1) + P(X=1)P(Y=0) \\
&= (1-p)p + p(1-p) \\
&= 2p(1-p)
\end{aligned}$$

$$\begin{aligned}
P(Z=2) &= P(X+Y=2) \\
&= P(X=1, Y=1) \\
(\text{independence}) &= P(X=1)P(Y=1) \\
&= p \cdot p \\
&= p^2
\end{aligned}$$

We have shown

$$P(Z=k) = \begin{cases} (1-p)^2 & k=0 \\ 2p(1-p) & k=1 \\ p^2 & k=2 \end{cases}$$

$$\begin{aligned}
P(Z=k) &= {}^2C_k p^k (1-p)^{2-k} \quad k=0,1,2 \\
\Rightarrow Z &\sim \text{Binomial}(2, p)
\end{aligned}$$

Ex:- Compare above computation with how we found

$$\begin{aligned}
&P(k \text{ successes in } n \text{ trials}) \\
P(x_1, \dots, x_n) &= \begin{cases} \text{III} \\ \# \{j: x_j = 1\} & \{j: x_j = 0\} \\ p & (1-p) \end{cases} \\
&\quad \downarrow \\
&{}^nC_k p^k (1-p)^{n-k}
\end{aligned}$$

Sums of Random Variables

Let $X, Y \sim \text{Bernoulli}(p)$ be two independent random variables.

$Z = X + Y, Z \sim ?$.

Expectation of product

X and Y are independent, then $E[XY] =$

Covariance

The “covariance of X and Y ” is defined as

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])].$$

Covariance-Facts

$$\text{Cov}[X, Y] = \text{Cov}[Y, X];$$

$$\text{Cov}[X, aY + bZ] = a \cdot \text{Cov}[X, Y] + b \cdot \text{Cov}[X, Z];$$

$$\text{Cov}[aX + bY, Z] = a \cdot \text{Cov}[X, Z] + b \cdot \text{Cov}[Y, Z]; \text{ and}$$

If X and Y are independent with a finite covariance, then

$$\text{Cov}[X, Y] = 0.$$

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$