

Independence

Introduction

Recall: Two events A, B are independent if $P(A \cap B) = P(A)P(B)$

We next want to extend the concept of independence to random variables

Definition: Random variables X and Y are independent if

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \quad \text{for any } A, B \subset \mathbb{R}.$$

Some Characterizations of Independence

1. **Theorem:** In general, random variables X and Y are independent iff

$$P(X \leq x, Y \leq y) = \underbrace{F(x, y)}_{\substack{\text{If} \\ \text{and only If}}} = F_X(x)F_Y(y) \quad \text{for any } x, y \in \mathbb{R}.$$

2. **Theorem:** Jointly discrete variables X and Y are independent iff

$$\underbrace{P(X = x, Y = y)}_{\substack{\text{Joint} \\ \text{Pmf}}} = f(x, y) = \underbrace{f_X(x)f_Y(y)}_{\substack{\text{Product of} \\ \text{marginal pmfs}}} = \underbrace{P(X = x)P(Y = y)}_{\substack{\text{Product of} \\ \text{marginal pmfs}}} \quad \text{for any } x, y \in \mathbb{R}.$$

3. **Theorem:** Jointly continuous variables X and Y are independent iff (there exist versions of the pdfs $f(x, y)$, $f_X(x)$ and $f_Y(y)$ such that)

$$f(x, y) = f_X(x)f_Y(y) \quad \text{for any } x, y \in \mathbb{R}.$$

(The “there exist versions” business is needed because one can change densities at a few values (i.e., endpoints) without changing the integrals computed from them.)

Independence

Introduction (cont'd)

4. Furthermore, if X and Y are independent so that $\underline{f(x,y)} = \underline{f_X(x)f_Y(y)}$ then

$$f(x|y) = f_X(x) \quad \text{for any } x \in \mathbb{R}, \text{ any } y \text{ with } \underline{f_Y(y)} > 0;$$

$$f(y|x) = f_Y(y) \quad \text{for any } y \in \mathbb{R}, \text{ any } x \text{ with } \underline{f_X(x)} > 0.$$

$$\rightarrow f_{x|y}(y) = \frac{f_{x,y}(x,y)}{f_Y(y)} \stackrel{\text{ind}}{=} \frac{\cancel{f_X(x)} f_Y(y)}{\cancel{f_Y(y)}} = f_X(x)$$

In other words, under independence, the conditional distribution of Y given $X = x$ is the same as the marginal distribution of Y . This is another way to think of independence: given information about one variable, the distribution of the 2nd variable does not change.

Two uses of independence

1. examine a joint pdf to determine if the r.v.s are independent
2. assume independence and derive joint distribution from marginal distributions

Independence

Some Examples

Discrete example:

		x	
	1	2	3
3	1/12	1/12	1/6
y	2	1/12	1/6
1	1/6	1/12	1/12

$\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

Are X and Y independent?

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

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

$$P(X=2, Y=1) = \frac{1}{12} \quad \text{(*)}$$

$$P(X=2) \cdot P(Y=1) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9} \quad \text{(**)}$$

$\text{(*)} \neq \text{(**)}$ X and Y are NOT independent.

Continuous example (the familiar one): $f(x,y) = 1/x$, $0 < y < x < 1$.

Are X and Y independent?

$$\rightarrow Y | X \sim \text{Uni}(0,x) \Rightarrow f(y|x) = \frac{1}{x} \quad \checkmark \quad 0 < y < x$$

$$f_Y(y) = \begin{cases} -\log y & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases} \quad f_Y(y) = -\log y \quad 0 < y < 1$$

$$\Rightarrow f(y|x) \neq f_Y(y) \Rightarrow$$

X and Y are NOT independent.

Another continuous example: $f(x,y) = 2y$, $0 \leq x \leq 1$, $0 \leq y \leq 1$.

Are X and Y independent?

$$f_X(x) = \int_0^1 f_{X,Y}(x,y) dy = \int_0^1 2y dy = y^2 \Big|_0^1 = 1 \quad \text{I}$$

$$\Rightarrow f_X(x) = \begin{cases} 0 & \text{If } x \notin (0,1) \\ 1 & \text{If } x \in (0,1) \end{cases}$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y) dx = \int_0^1 2y dx = 2y \Big|_0^1 = 2y \quad \text{II}$$

$$f_Y(y) = \begin{cases} 2y & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

X and Y are independent.

$$2y = f_{X,Y}(x,y) = f_X(x) f_Y(y) = 1 \cdot 2y$$

$$f_{X,Y}(x,y) \in (0,1) \times (0,1)$$

$g(X)$ is R.V.

$$g(X): \Omega \xrightarrow{\quad} \mathbb{R}$$

Independence

Some important results & facts



1. Theorem: If X, Y are independent then $g(X)$ and $h(Y)$ are independent.

Proof: $P(g(X) \in A \text{ and } h(Y) \in B)$

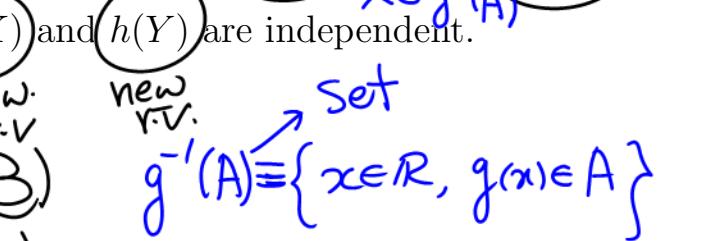
$$= P(X \in g^{-1}(A) \text{ and } Y \in h^{-1}(B))$$

X and Y are ind.

$$= P(X \in g^{-1}(A)) P(Y \in h^{-1}(B))$$

✓

$$= P(g(X) \in A) \cdot P(h(Y) \in B)$$



Ex: X and Y are independent then e^X and $\sin Y$ are independent.

2. Theorem: If X, Y are independent and the mean of $g(X)h(Y)$ exists, then

$$\rightarrow E[g(X)h(Y)] = Eg(X)Eh(Y) \quad \boxed{E(XY) = E(X)E(Y)}$$

$$E(Z(X,Y)) \stackrel{\text{def}}{=} \iint Z(x,y) f_{XY}(x,y) dx dy = \iint g(x)h(y) f_X(x) f_Y(y) dx dy$$

$$= \int g(w) f_X(w) dw \int h(y) f_Y(y) dy = E[g(X)] E[h(Y)].$$

3. Corollary: If X, Y are independent and $E(XY)$ exists, then

$$\text{Proof: } \text{Cov}(X, Y) \stackrel{\text{def}}{=} E[XY] - E[X]E[Y] \quad \frac{\text{X and Y are ind.}}{\text{X and Y are ind.}} \quad E[X]E[Y] - E[X]E[Y] = 0$$

So independent $\Rightarrow 0$ correlation. But, it is easy enough to find examples showing that 0 correlation does NOT imply independence.

Independence

Some important results & facts

4. Corollary: If X, Y are independent and $\mathbb{E}X^2, \mathbb{E}Y^2$ exist, then

$$\text{Var}(aX + bY) = \text{Var}(aX) + \text{Var}(bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$$

$$\text{Var}(aX + bY) = \text{Var}(aX) + \text{Var}(bY) + 2ab \underbrace{\text{Cov}(X, Y)}_{=0}$$

In general

$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i) \quad \text{If } X_1, X_2, \dots, X_n \text{ are independent}$$

- * 5. Corollary: If X, Y are independent and X has mgf $M_X(t)$ and Y has $M_Y(t)$ then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

holds for any t where $M_X(t)$ and $M_Y(t)$ exist.

(Note $M_X(t)$ must exist in some $(-h_1, h_1)$ and $M_Y(t)$ must exist in some $(-h_2, h_2)$, so that both $M_X(t)$ and $M_Y(t)$ exist in some neighborhood of zero.)

$$M_{X+Y}(t) \stackrel{\text{def}}{=} \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] \quad \begin{matrix} X \neq Y \text{ are} \\ \text{independent} \end{matrix}$$

$$\mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = M_X(t) M_Y(t).$$

$$M_{X_1+X_2+\dots+X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_n}(t) \quad \text{If } X_1, X_2, \dots, X_n \text{ are independent}$$

6. Theorem: Suppose X has mgf $M_X(t)$ and Y has $M_Y(t)$. Then, X, Y are independent iff

$$M_{X,Y}(t_1, t_2) = \mathbb{E}e^{t_1 X + t_2 Y} = \mathbb{E}e^{t_1 X} \mathbb{E}e^{t_2 Y} = M_X(t_1)M_Y(t_2) \quad \begin{matrix} \text{two arguments} \\ t_1, t_2 \end{matrix}$$

holds for all $t_1, t_2 \in (-h, h)$ in some neighborhood of zero.

$$M_{X,Y}(t_1, t_2) = \frac{1}{(1-t_1)(1-t_1-t_2)} \quad \begin{matrix} t_1 < 1 \\ t_1+t_2 < 1 \end{matrix}$$

$$M_X(t_1) = M_{X,Y}(t_1, 0) = \frac{1}{(1-t_1)^2} \implies X \text{ and } Y \text{ are NOT independent.}$$

$$M_Y(t_2) = M_{X,Y}(0, t_2) = \frac{1}{(1-t_2)^2} \quad M_X(t_1) = e^{t_1^2/2} \quad M_Y(t_2) = e^{t_2^2/2}$$

Independence

Multivariate case

Definition: Random variables X_1, \dots, X_n are independent if

$$\rightarrow P(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^n P(X_i \in A_i) \quad \text{for any } A_1, \dots, A_n \subset \mathbb{R}$$

Characterizations

1. X_1, \dots, X_n are **independent** iff

$$\underbrace{F(x_1, \dots, x_n)}_{\substack{\\ \text{for any } x_1, \dots, x_n \in \mathbb{R}}} = P(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^n F_{X_i}(x_i)$$

2. X_1, \dots, X_n are **independent** iff

$$\underbrace{f(x_1, \dots, x_n)}_{\substack{\\ \text{for any } x_1, \dots, x_n \in \mathbb{R}}} = \prod_{i=1}^n f_{X_i}(x_i)$$



Important Properties: If X_1, \dots, X_n are independent r.v.s, then

1. $g_1(X_1), \dots, g_n(X_n)$ are independent r.v.s

2. it holds that $E \left[\prod_{i=1}^n g_i(X_i) \right] = \prod_{i=1}^n E[g_i(X_i)]$

3. $\text{Cov}(X_i, X_j) = 0$ for $i \neq j$

4. $\text{Var} \left(\sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i)$

5. if the mgf $M_{X_i}(t)$ exists for each $i = 1, \dots, n$, then the mgf of $X_1 + \dots + X_n$ is

$$M_{\underbrace{X_1 + \dots + X_n}_{\substack{\\ \text{for all } t \text{ where each } M_{X_i}(t) \text{ exists.}}}}(t) = \prod_{i=1}^n M_{X_i}(t)$$

6. MGF characterization: X_1, \dots, X_n are independent iff

$$M_{X_1, \dots, X_n}(t_1, \dots, t_n) = \prod_{i=1}^n M_{X_i}(t_i)$$