Random variables

March 6, 2019

Intro

- Sample space and events are basic components of probability
- Similar to numbers in calculus
- Study the relations between numbers we use functions
- In probability we use *random variables*

Intro

- Sample space and events are basic components of probability
- Similar to numbers in calculus
- Study the relations between numbers we use functions
- In probability we use *random variables*

Intro

- Sample space and events are basic components of probability
- Similar to numbers in calculus
- Study the relations between numbers we use functions
- In probability we use *random variables*

Intro

- Sample space and events are basic components of probability
- Similar to numbers in calculus
- Study the relations between numbers we use functions
- In probability we use *random variables*

Definition

Let (Ω, P) be a probability space.
A **random variable** X defined on (Ω, P) is a quantity that is calculated by the outcomes in Ω.

- Can think of random variable (RV) as function of outcomes
- RV can take different values with different outcomes
- Can calculate probability of each value of RV

- Can think of random variable (RV) as function of outcomes
- RV can take different values with different outcomes
- Can calculate probability of each value of RV

- Can think of random variable (RV) as function of outcomes
- RV can take different values with different outcomes
- Can calculate probability of each value of RV

Example

Roll 2 dice. X= sum of 2 dice

- X takes values from 2 to 12
- Can calculate

$$P(X = j), \;\; j = 2, 3, \ldots, 12$$

Example

Roll 2 dice. X= sum of 2 dice

- X takes values from 2 to 12
- Can calculate

$$P(X=j), \quad j=2,3,\ldots,12$$

Distribution

- $P(X = 2) = P(X = 12) = 1/36$
- $P(X = 3) = P(X = 11) = 2/36$
- $P(X = 4) = P(X = 10) = 3/36$
- $P(X = 5) = P(X = 9) = 4/36$
- $P(X = 6) = P(X = 8) = 5/36$
- $P(X = 7) = 6/36$

Distribution

- X has its own set of all possible values called the range of X
- For each value x there is a probability that $X = x$.
- X creates a new probability space with
$$\sum_x P(X = x) = 1$$
- This probability is called the distribution of X

Example

Suppose that an individual purchases two electronic components each of which may be either defective or acceptable. Suppose that the four possible results - $(d, d), (d, a), (a, d), (a, a)$ - have respective probabilities .09, .21, .21, .49 [d means defective, a means acceptable, and so on]. Let X denote the number of acceptable components obtained in the purchase. Find the distribution of X.

Solution

- Values of X: 0,1,2
- $P(X = 0) = .09$
- $P(X = 1) = .21 + .21 = .42$
- $P(X = 2) = .49$

Solution

- Values of X: 0,1,2
- $P(X = 0) = .09$
- $P(X = 1) = .21 + .21 = .42$
- $P(X = 2) = .49$

Solution

- Values of X: 0,1,2
- $P(X = 0) = .09$
- $P(X = 1) = .21 + .21 = .42$
- $P(X = 2) = .49$

Solution

- Values of X: 0,1,2
- $P(X = 0) = .09$
- $P(X = 1) = .21 + .21 = .42$
- $P(X = 2) = .49$

Example

- Same example, want to know if there is at least one acceptable component?
- Define new RV

$$I = \begin{cases} 1 & \text{if } X = 1 \text{ or } 2 \\ 0 & \text{if } X = 0 \end{cases}$$

- I : the *indicator RV* of the event "at least one acceptable"

Example

- Same example, want to know if there is at least one acceptable component?
- Define new RV

$$I = \begin{cases} 1 & \text{if } X = 1 \text{ or } 2 \\ 0 & \text{if } X = 0 \end{cases}$$

- I : the *indicator RV* of the event "at least one acceptable"

Example

$$P(\text{at least one acceptable})$$
$$= P(I = 1) = 0.91$$

Types of RV

- If the set of values of X $Range(X)$ is finite or countable then X is called *discrete RV*
- If the set of values of X is uncountable (like the interval $[a, b]$) then X is called *continuous RV*

Types of RV

- If the set of values of X $Range(X)$ is finite or countable then X is called *discrete RV*
- If the set of values of X is uncountable (like the interval $[a, b]$) then X is called *continuous RV*

Distribution function

The **cumulative distribution function** (cdf) F of a random variable X is defined as

$$F(x) = P(X \leq x)$$

for any real value x.

Each RV has only one cdf.

Distribution function

The **cumulative distribution function** (cdf) F of a random variable X is defined as

$$F(x) = P(X \leq x)$$

for any real value x.
Each RV has only one cdf.

Example

X is the number of acceptable components as in previous example.

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ .09 & \text{if } 0 \le x < 1 \\ .51 & \text{if } 1 \le x < 2 \\ 1 & \text{if } 2 \le x \end{cases}$$

Example

X is the number of acceptable components as in previous example.

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ .09 & \text{if } 0 \leq x < 1 \\ .51 & \text{if } 1 \leq x < 2 \\ 1 & \text{if } 2 \leq x \end{cases}$$

Using cdf

$$P(a < X \leq b) = F(b) - F(a)$$

Example

Suppose X has cdf

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 - e^{-x^2} & \text{if } x > 0 \end{cases}$$

What is the probability of $X > 1$?

Solution

$$
\begin{aligned}
P(X > 1) &= 1 - P(X \leq 1) \\
&= 1 - F(1) \\
&= e^{-1} \\
&= .368
\end{aligned}
$$

Probability mass function

Let X be a discrete RV. The *probability mass function* (pmf) of X is defined as

$$p(x) = P(X = x) \text{ for all } x \in \Omega$$

Properties

- X is discrete
  $\rightarrow Range(X) = \{x_1, \dots, x_n \dots\}$
- $p(x_i) = P(X = x_i) \geq 0$
- $\sum_{i=1}^{\infty} p(x_i) = 1$
- cdf: $F(x) = \sum_{x_i < x} p(x)$

Properties

- X is discrete
  $\rightarrow Range(X) = \{x_1, \ldots, x_n \ldots\}$
- $p(x_i) = P(X = x_i) \geq 0$
- $\sum_{i=1}^{\infty} p(x_i) = 1$
- cdf: $F(x) = \sum_{x_i < x} p(x)$

Properties

- X is discrete
  $\rightarrow Range(X) = \{x_1, \ldots, x_n \ldots\}$
- $p(x_i) = P(X = x_i) \geq 0$
- $\sum_{i=1}^{\infty} p(x_i) = 1$
- cdf: $F(x) = \sum_{x_i < x} p(x)$

Properties

- X is discrete
  $\rightarrow Range(X) = \{x_1, \ldots, x_n \ldots\}$
- $p(x_i) = P(X = x_i) \geq 0$
- $\sum_{i=1}^{\infty} p(x_i) = 1$
- cdf: $F(x) = \sum_{x_i < x} p(x)$

Example

Suppose X has 3 values $1, 2, 3$ and

$$p(1) = \frac{1}{2}, \; p(2) = \frac{1}{3}$$

then what is $p(3)$?

$$p(3) = 1 - p(1) - p(2) = 1/6.$$

Example

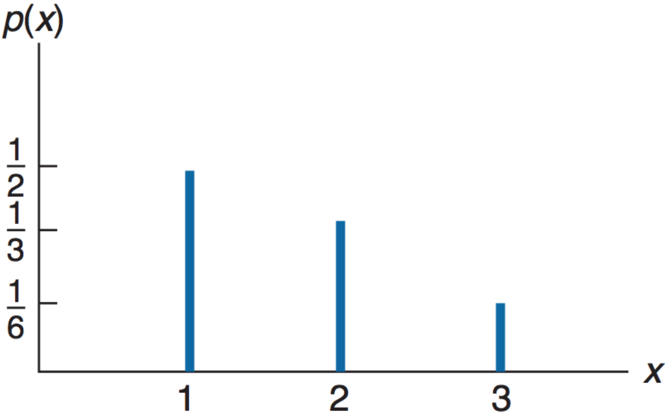
Suppose X has 3 values $1, 2, 3$ and

$$p(1) = \frac{1}{2},\ p(2) = \frac{1}{3}$$

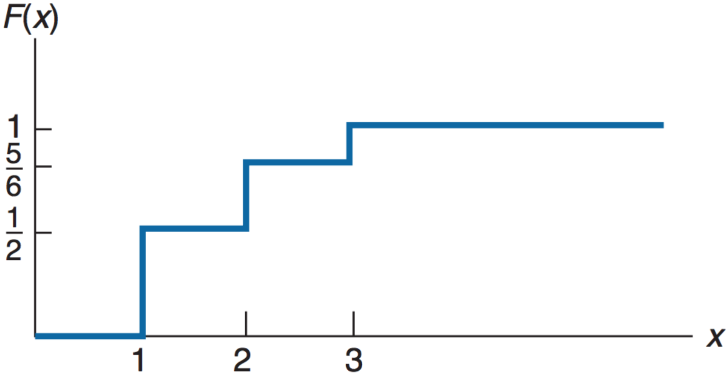
then what is $p(3)$?

$$p(3) = 1 - p(1) - p(2) = 1/6.$$

Graph of $p(x)$



Graph of $F(x)$



Probability density functions

Suppose $Range(X)$ is uncountable. X is *continuous* if there is a non negative function $f(x)$ so that

$$P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

$f(x)$ is called the *probability density function* of X.
If $F(x)$ is differentiable then $f(x) = F'(x)$

Properties

- $1 = \int\limits_{-\infty}^{\infty} f(x)dx$
- $P(a < X < b) = \int\limits_{a}^{b} f(x)dx$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$
- $P(X = a) = \int_a^a f(x)dx = 0$

Properties

- $1 = \int_{-\infty}^{\infty} f(x)dx$
- $P(a < X < b) = \int_{a}^{b} f(x)dx$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$
- $P(X = a) = \int_a^a f(x)dx = 0$

Properties

- $1 = \int\limits_{-\infty}^{\infty} f(x)dx$
- $P(a < X < b) = \int\limits_{a}^{b} f(x)dx$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$
- $P(X = a) = \int_a^a f(x)dx = 0$

Properties

- $1 = \int\limits_{-\infty}^{\infty} f(x)dx$
- $P(a < X < b) = \int\limits_{a}^{b} f(x)dx$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$
- $P(X = a) = \int_a^a f(x)dx = 0$

Example

Suppose the pdf of X is

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of C?
Find $P(X > 1)$?

Solution

By property of pdf:
$\int_0^2 C(4x - 2x^2)dx = 1$, which
implies $C = 3/8$.

Then

$$P(X > 1) = \int_1^2 \frac{3}{8}(4x - 2x^2)dx = \frac{1}{2}$$

Solution

By property of pdf:
$\int_0^2 C(4x-2x^2)dx = 1$, which implies $C = 3/8$.
Then

$$P(X>1) = \int_1^2 \frac{3}{8}(4x-2x^2)dx = \frac{1}{2}$$

Practice

Let X represent the difference between the number of heads and the number of tails obtained when a coin is tossed n times. What are the possible values of X ?

Practice

7. The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \leq 100 \\ \dfrac{100}{x^2} & x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? Assume that the events $E_i, i = 1, 2, 3, 4, 5$, that the ith such tube will have to be replaced within this time are independent.

Jointly distributed random variables

Joint cdf

- Suppose X and Y are 2 RVs based on the same experiment.
- We wish to study the relation between them.
- Joint cumulative distribution function

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

Joint cdf

- Suppose X and Y are 2 RVs based on the same experiment.
- We wish to study the relation between them.
- Joint cumulative distribution function

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

cdf from joint cdf

- $P(X \leq x) = F_X(x)$
  $= F_{X,Y}(X \leq x, Y \leq \infty)$
- $P(Y \leq y) = F_Y(y)$
  $= F_{X,Y}(X \leq \infty, Y \leq y)$

cdf from joint cdf

- $P(X \leq x) = F_X(x)$
  $= F_{X,Y}(X \leq x, Y \leq \infty)$
- $P(Y \leq y) = F_Y(y)$
  $= F_{X,Y}(X \leq \infty, Y \leq y)$

Joint pmf

X and Y are discrete: joint pmf

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

pmf from joint pmf

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

Example

Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries. If we let X and Y denote, respectively, the number of new and used but still working batteries that are chosen, find the joint probability mass function of X and Y

Solution

TABLE 4.1 $P\{X = i, Y = j\}$

i \ j	0	1	2	3	Row Sum $= P\{X = i\}$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
Column Sums $= P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

Joint pdf

X and Y are continuous: joint pdf $f_{X,Y}(x,y)$ so that

$$F_{X,Y}(a,b) = \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x,y)dxdy$$

and

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

Joint pdf

X and Y are continuous: joint pdf $f_{X,Y}(x,y)$ so that

$$F_{X,Y}(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x,y)dxdy$$

and

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

Example

X and Y have joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{if } 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X < Y)$.

Solution

$$
\begin{aligned}
P\{X < Y\} &= \iint\limits_{(x,y):x<y} 2e^{-x}e^{-2y}\,dx\,dy \\
&= \int_0^\infty \int_0^y 2e^{-x}e^{-2y}\,dx\,dy \\
&= \int_0^\infty 2e^{-2y}(1 - e^{-y})\,dy \\
&= \int_0^\infty 2e^{-2y}\,dy - \int_0^\infty 2e^{-3y}\,dy \\
&= 1 - \frac{2}{3} \\
&= \frac{1}{3}
\end{aligned}
$$

pdf from joint pdf

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

$$f_Y(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

Practice

10. The joint probability density function of X and Y is given by

$$f(x, y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right), \quad 0 < x < 1, \quad 0 < y < 2$$

(a) Verify that this is indeed a joint density function.
(b) Compute the density function of X.
(c) Find $P\{X > Y\}$.

Independence

X and Y are independent if for any x, y

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

Can check independency by looking at cdf, pmf and pdf

If X and Y are independent then

- $F_{X,Y}(x,y) = F_X(x)F_Y(y)$
- $p_{X,Y}(x,y) = p_X(x)p_Y(y)$
- $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

If X and Y are independent then

- $F_{X,Y}(x, y) = F_X(x)F_Y(y)$
- $p_{X,Y}(x, y) = p_X(x)p_Y(y)$
- $f_{X,Y}(x, y) = f_X(x)f_Y(y)$

If X and Y are independent then

- $F_{X,Y}(x, y) = F_X(x)F_Y(y)$
- $p_{X,Y}(x, y) = p_X(x)p_Y(y)$
- $f_{X,Y}(x, y) = f_X(x)f_Y(y)$

Practice

A bin of 5 transistors is known to contain 3 that are defective.The transistors are to be tested, one at a time, until the defective ones are identified. Denote by N_1 the number of tests made until the first defective is spotted and by N_2 the number of additional tests until the second defective is spotted; find the joint probability mass function of N_1 and N_2.

Homework 4

Chapter 4: 1, 3, 4, 6, 8, 10, 12