STA286 Lecture 09

Neil Montgomery

Last edited: 2017-02-01 10:05

extension to more than two random variables - a few illustrations

The concepts of joint and marginal distributions extend to n random variables at a time. Things get ugly fast. Here is a taste. . . .

extension to more than two random variables - a few illustrations

The concepts of joint and marginal distributions extend to n random variables at a time. Things get ugly fast. Here is a taste. . . .

Joint pmf of X_1, X_2, X_3:

$$p(x_1, x_2, x_3) = P(X_1 = x_1, X_2 = x_2, X_3 = x_3)$$

extension to more than two random variables - a few illustrations

The concepts of joint and marginal distributions extend to n random variables at a time. Things get ugly fast. Here is a taste....

Joint pmf of X_1, X_2, X_3:

$$p(x_1, x_2, x_3) = P(X_1 = x_1, X_2 = x_2, X_3 = x_3)$$

Marginal for X_2:

$$p(x_2) = \sum_{x_1} \sum_{x_3} p(x_1, x_2, x_3)$$

extension to more than two random variables - a few illustrations

The concepts of joint and marginal distributions extend to n random variables at a time. Things get ugly fast. Here is a taste. . . .

Joint pmf of X_1, X_2, X_3:

$$p(x_1, x_2, x_3) = P(X_1 = x_1, X_2 = x_2, X_3 = x_3)$$

Marginal for X_2:

$$p(x_2) = \sum_{x_1} \sum_{x_3} p(x_1, x_2, x_3)$$

Joint density for Y_1, Y_2, Y_3:

$$P(a_1 < Y_1 < b_1, a_2 < Y_2 < b_2, a_3 < Y_3 < b_3) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(y_1, y_2, y_3)\, dy_3\, dy_2\, dy_1$$

extension to more than two random variables - a few illustrations

The marginal density for X_3:

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, x_3)\, dx_1\, dx_2$$

extension to more than two random variables - a few illustrations

The marginal density for X_3:

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2, x_3)\, dx_1\, dx_2$$

Important: Random variables $X_1, X_2, \ldots, X_n$ are *independent* if:

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$

(continuous version with densities. . . discrete version is similar.)

extension to more than two random variables - a few illustrations

The marginal density for X_3:

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, x_3)\, dx_1\, dx_2$$

Important: Random variables $X_1, X_2, \ldots, X_n$ are *independent* if:

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$

(continuous version with densities. . . discrete version is similar.)

This is conceptually important as the basis for the mathematical model of the observations in a dataset.

the expected value operator

the mean of a distribution

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

the mean of a distribution

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

A (discrete) random variable can have an "average", which is a weighted sum of the outcomes with their probabilities as weights.

the mean of a distribution

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

A (discrete) random variable can have an "average", which is a weighted sum of the outcomes with their probabilities as weights.

BIG MONEY. We play a gambling game called BIG MONEY. Roll a die. This is your outcome after each play of BIG MONEY:

Roll	1, 2	3	4, 5	6
Outcome $	-2	0	1	2

the mean of a distribution

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

A (discrete) random variable can have an "average", which is a weighted sum of the outcomes with their probabilities as weights.

BIG MONEY. We play a gambling game called BIG MONEY. Roll a die. This is your outcome after each play of BIG MONEY:

Roll	1, 2	3	4, 5	6
Outcome \$	-2	0	1	2
Probability	2/6	1/6	2/6	1/6

BIG MONEY is a "fair game"

Your *expected* financial outcome is (theoretically):

$$-2\frac{2}{6} + 0\frac{1}{6} + 1\frac{2}{6} + 2\frac{1}{6} = 0$$

BIG MONEY is a "fair game"

Your *expected* financial outcome is (theoretically):

$$-2\frac{2}{6}+0\frac{1}{6}+1\frac{2}{6}+2\frac{1}{6}=0$$

The *expected value* of a discrete random variable X is:

$$E(X)=\sum_{x} xp(x)$$

if the sum exists (actually it has to converge absolutely.)

expected number of coin tosses to first $\mathtt{H}$

Denote by X the number of coin tosses until the first $\mathtt{H}$. The pmf is (new version!):

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is (new version!):

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$E(X) = \sum_{x=1}^{\infty} x(1-p)^{x-1}p$$

expected number of coin tosses to first $\mathtt{H}$

Denote by X the number of coin tosses until the first $\mathtt{H}$. The pmf is (new version!):

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x(1-p)^{x-1} p \\ &= p \sum_{x=0}^{\infty} x(1-p)^{x-1} \end{aligned}$$

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is (new version!):

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x(1-p)^{x-1} p \\ &= p \sum_{x=0}^{\infty} x(1-p)^{x-1} \\ &= p \sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^{x} \end{aligned}$$

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is (new version!):

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} x(1-p)^{x-1} p \\
&= p \sum_{x=0}^{\infty} x(1-p)^{x-1} \\
&= p \sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^{x} \\
&= p \frac{d}{dp}\left(\sum_{x=0}^{\infty} (1-p)^{x}\right) = p \frac{d}{dp}\left(\frac{1}{1-(1-p)}\right) = \frac{1}{p}
\end{aligned}$$

generalization of "tosses to first head"

A factory makes a defective item with probability p (per item) with $0 < p < 1$. What is the expected number of items until the first defective item?

Denote the number of items by X. The pmf of X will be:

$$p(x) = (1-p)^{x-1}\, p$$

for $x \in \{1, 2, 3, \ldots\}$.

generalization of “tosses to first head”

A factory makes a defective item with probability p (per item) with $0 < p < 1$. What is the expected number of items until the first defective item?

Denote the number of items by X. The pmf of X will be:

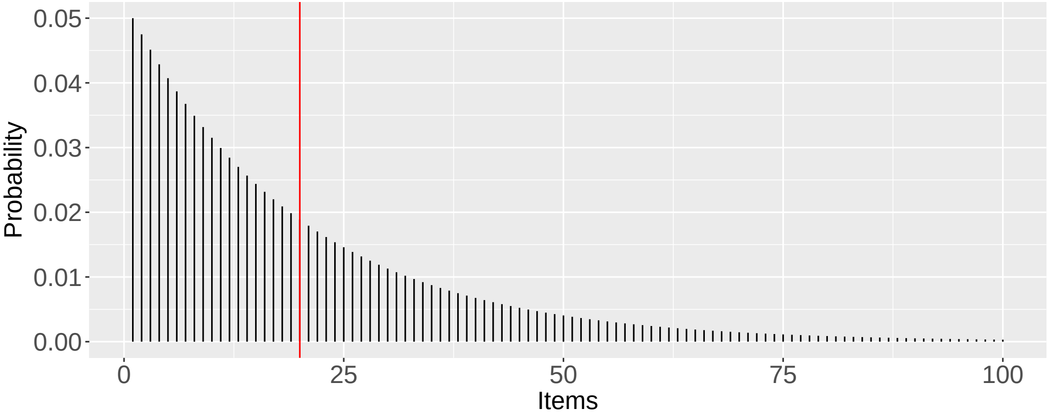
$$p(x) = (1 - p)^{x-1}\, p$$

for $x \in \{1, 2, 3, \ldots\}$.

According to the previous slide, $E(X) = \frac{1}{p}$

graphical view of expected value

Suppose $p = 0.05$. Then $E(X) = 20$. The expected value is the "physical" balance point of the pmf (a.k.a. the *first moment*)



expected value non-example

Let X have the following pmf (!):

$$p(x) = \frac{6}{\pi^2 x^2}, x \in \{1, 2, 3, \ldots\}$$

X does not have an expected value, because $\sum_x xp(x)$ does not converge.

expected value - continuous version

If X is continuous with density f, its expected value is:

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

provided the integral converges (absolutely).

expected value - continuous version

If X is continuous with density f, its expected value is:

$$E(X) = \int_{-\infty}^{\infty} x f(x) \, dx$$

provided the integral converges (absolutely).

Bus stop example. What is the expected waiting time?

expected value - continuous version

If X is continuous with density f, its expected value is:

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$$

provided the integral converges (absolutely).

Bus stop example. What is the expected waiting time?

$$\int_0^{10} x\frac{1}{10}\,dx = x^2\frac{1}{20}\bigg|_{x=0}^{10} = 5$$

another continuous example; plus a non-example

Consider $f(x) = 2x^{-3}$ for $x > 1$, and 0 otherwise. This is a valid density.

another continuous example; plus a non-example

Consider $f(x) = 2x^{-3}$ for $x > 1$, and 0 otherwise. This is a valid density.

Suppose X has density $f(x)$. Then:

$$E(X) = \int_1^\infty x\, 2x^{-3}\, dy$$

another continuous example; plus a non-example

Consider $f(x) = 2x^{-3}$ for $x > 1$, and 0 otherwise. This is a valid density.

Suppose X has density $f(x)$. Then:

$$E(X) = \int_1^\infty x\, 2x^{-3}\, dy = \left[-2x^{-1}\right]_{x=1}^\infty$$

another continuous example; plus a non-example

Consider $f(x) = 2x^{-3}$ for $x > 1$, and 0 otherwise. This is a valid density.

Suppose X has density $f(x)$. Then:

$$E(X) = \int_1^\infty x\, 2x^{-3}\, dy = \left[-2x^{-1}\right]_{x=1}^{\infty} = 2$$

another continuous example; plus a non-example

Consider $f(x) = 2x^{-3}$ for $x > 1$, and 0 otherwise. This is a valid density.

Suppose X has density $f(x)$. Then:

$$E(X) = \int_1^\infty x\, 2x^{-3}\, dy = \left[-2x^{-1}\right]_{x=1}^\infty = 2$$

But suppose Y has density $f(y) = y^{-2}$ on $y > 1$, and 0 otherwise. Then:

$$\int_1^\infty y\, y^{-2}\, dy = [\log y]_{y=1}^\infty$$

so $E(Y)$ does not exist.

expected values of functions of random variables

Denote by X your outcome after a play of BIG MONEY. Then...

expected values of functions of random variables

Denote by X your outcome after a play of BIG MONEY. Then. . .

BIG MONEY—*after a few schnapps version*.

expected values of functions of random variables

Denote by X your outcome after a play of BIG MONEY. Then...

BIG MONEY—*after a few schnapps version*.

We adopt fake German accents and up the game.

Roll	1, 2	3	4, 5	6
Outcome $	-200	0	100	200
Probability	2/6	1/6	2/6	1/6

expected values of functions of random variables

Denote by X your outcome after a play of BIG MONEY. Then. . .

BIG MONEY—*after a few schnapps version*.

We adopt fake German accents and up the game.

Roll	1, 2	3	4, 5	6
Outcome $	-200	0	100	200
Probability	2/6	1/6	2/6	1/6

Denote by Y your outcome after a play of this modified game. $E(Y)$ is also zero, using the definition:

$$E(Y) = -200\frac{2}{6} + 0\frac{1}{6} + 100\frac{2}{6} + 200\frac{1}{6} = 0$$

expected values of functions of random variables

The definition requires determining the pmf/pdf of the new random variable, which is not always so easy.

expected values of functions of random variables

The definition requires determining the pmf/pdf of the new random variable, which is not always so easy.

Theorem: For a random variable X and a function $g : \mathbb{R} \to \mathbb{R}$:

$$E(g(X)) = \sum_{x} g(x)p(x) \qquad \text{discrete} \qquad E(g(X)) = \int_{\infty}^{\infty} g(x)f(x)\,dx \qquad \text{continuous}$$

expected values of functions of random variables

The definition requires determining the pmf/pdf of the new random variable, which is not always so easy.

Theorem: For a random variable X and a function $g : \mathbb{R} \to \mathbb{R}$:

$$E(g(X)) = \sum_x g(x)p(x) \qquad \text{discrete} \qquad E(g(X)) = \int_{\infty}^{\infty} g(x)f(x)\,dx \qquad \text{continuous}$$

For example, the outcome Y of BIG MONEY (SCHNAPPS VERSION) is related the the outcome X of BIG MONEY by:

$$Y = 100X$$

So the theorem says $E(Y) = E(100X)$ and the calculation is (technically):

$$E(Y) = 100 \cdot (-2)\frac{2}{6} + 100 \cdot (0)\frac{1}{6} + 100 \cdot (1)\frac{2}{6} + 100 \cdot (2)\frac{1}{6} = 0$$

$E(\cdot)$ rules

The theorem lets us develop some basic rules.

$$\begin{aligned} E(a + bX) &= \sum_x (a + bx)p(x) \\ &= \sum_x ap(x) + \sum_x bxp(x) \end{aligned}$$

$E(\cdot)$ rules

The theorem lets us develop some basic rules.

$$\begin{aligned}
E(a + bX) &= \sum_x (a + bx)p(x) \\
&= \sum_x ap(x) + \sum_x bxp(x) \\
&= a\sum_x p(x) + b\sum_x xp(x)
\end{aligned}$$

$E(\cdot)$ rules

The theorem lets us develop some basic rules.

$$
\begin{aligned}
E(a+bX) &= \sum_x (a+bx)p(x) \\
&= \sum_x ap(x) + \sum_x bxp(x) \\
&= a\sum_x p(x) + b\sum_x xp(x) \\
&= a + bE(X)
\end{aligned}
$$

$E(\cdot)$ rules

The theorem lets us develop some basic rules.

$$\begin{aligned} E(a+bX) &= \sum_x (a+bx)p(x) \\ &= \sum_x ap(x) + \sum_x bxp(x) \\ &= a\sum_x p(x) + b\sum_x xp(x) \\ &= a + bE(X) \end{aligned}$$

Continuous version is the same.