Multiple Random Variables

David Gerard

2017-09-28

Learning Objectives

- Joint/conditional/marginal distributions of random variables.
- Independence of random variables.
- Covariance of random variables.
- Mean and variance of sum of random variables.

Motivating Example

- Consider experiment of tossing two fair dice.
- Sample space $= S = \{(1,1),\ (1,2),\ (1,3), \ldots,\ (5,6),\ (6,6)\}$.
- $P(z) = 1/36$ for all $z \in S$.
- Let $X =$ sum of the two dice.
- Let $Y = |\text{difference of the two dice}|$.
- For each of the 36 outcomes in the sample space, we have corresponding values of X and Y.
- We can write out the joint distribution probability mass function of X and Y.

joint pmf

joint pmf

Let X and Y be two discrete bivariate random variables. Then the function $f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ defined by $P(X = x \text{ and } Y = y)$ is called the joint probability mass function of X and Y.

Dice Example

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
y	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

$f(x,y) = 0$ for all other (x,y).

E.g.

$$
\begin{aligned}
f(3,1) &= P(X=3, Y=1) = P(\{(2,1),(1,2)\}) \\
&= P((2,1)) + P((1,2)) \text{ (disjoint events)} \\
&= \frac{1}{36} + \frac{1}{36} = \frac{1}{18}.
\end{aligned}
$$

Property of joint pmf

Some properties readily generalize to the multiple variable case:

- $\sum_x \sum_y f(x, y) = 1$
- $0 \leq f(x, y) \leq 1$ for all (x, y).

Marginal Distributions

From the joint pmf, we can find the pmf of **just** X and the pmf of **just** Y.

To find the pmf of X at a particular x, we find every pair (x, y) and add up their joint probabilities.

$$f_X(x) = \sum_y f(x, y) = \sum_y P(X = x \text{ and } Y = y).$$

Similarly

$$f_Y(y) = \sum_x f(x, y) = \sum_x P(X = x \text{ and } Y = y).$$

Even though f_X and f_Y are just the pmf's of X and Y, we call them marginal pmfs when in the context of multiple variables.

Dice Example

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
y	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

$$P(X=6) = f_x(6) = f(6,0) + f(6,3) + f(6,4)$$
$$= 1/36 + 1/18 + 1/18 = 5/36$$

Dice Example

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
y	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

$$
\begin{aligned}
P(Y=2) = f_y(2) &= f(4,2) + f(6,2) + f(8,2) + f(10,2) \\
&= 1/18 + 1/18 + 1/18 + 1/18 = 4/36
\end{aligned}
$$

Conditional Distribution

Conditional pmf

Let X and Y be discrete random variables. For any x such that $P(X = x) = f_X(x) > 0$, the conditional pmf of Y given $X = x$ is

$$f_{Y|X}(y|x) = \frac{P(X = x \text{ and } Y = y)}{P(X = x)} = \frac{f(x, y)}{f_X(x)}.$$

Similarly

$$f_{X|Y}(x|y) = \frac{P(X = x \text{ and } Y = y)}{P(Y = y)} = \frac{f(x, y)}{f_Y(y)}.$$

Dice Example i: $f(x|y=2)$

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
y	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

Dice Example ii: $f(x|y=2)$

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
	0											
	1											
y	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3											
	4											
	5											

Intuition: Normalize selected row to sum to 1 so that they are probabilities.

Dice Example ii: $f(x|y=2)$

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
	0											
	1											
y	2			$\frac{1}{4}$		$\frac{1}{4}$		$\frac{1}{4}$		$\frac{1}{4}$		
	3											
	4											
	5											

$$\frac{1}{18}+\frac{1}{18}+\frac{1}{18}+\frac{1}{18}=4/18$$

$$\frac{1/18}{4/18}=\frac{1}{4}$$

$f(x,y)$		x										
		2	3	4	5	6	7	8	9	10	11	12
	0											
	1											
y	2			$\frac{1}{4}$		$\frac{1}{4}$		$\frac{1}{4}$		$\frac{1}{4}$		
	3											
	4											
	5											

So

$$f_{X|Y}(x|y=2) = \begin{cases} \frac{1}{4} & \text{if } x \in \{4,6,8,10\} \\ 0 & \text{otherwise.} \end{cases}$$

Independence

independence

Two random variables X and Y are independent if for every $x \in \mathbb{R}$ and $y \in \mathbb{R}$

$$f(x, y) = f_X(x)f_Y(y).$$

- This is equivalent to $f_{X|Y}(x|y) = f_X(x)$ and $f_{Y|X}(y|x) = f_Y(y)$ for all x and y.
- I.e. knowing y doesn't change the distribution of x and knowing x doesn't change the distribution of y.
- So knowing y doesn't tell you anything new about x and knowing x doesn't tell you anything new about y.

Extension of law of unconscious statistician

With two random variables X, Y we can create a new random variable $Z = g(X, Y)$ with $g : R \times R \to R$.

For example $g(x, y) = x + y$ or $g(x, y) = x \cdot y$.

Again we could try to compute the distribution of Z - hard (conscious statistician).

Or we have:

Law of Unconscious Statistician

$$E(g(X, Y)) = \sum_{x \in R_x, y \in R_y} g(x, y) f_{X,Y}(x, y)$$

Consequences of law of unconscious statistician i

$E(X+Y) = EX + EY.$

$$\begin{aligned} E(X+Y) &= \sum_{x\in R_x, y\in R_y} (x+y) f_{X,Y}(x,y) \\ &= \sum_{x\in R_x, y\in R_y} x f_{X,Y}(x,y) + \sum_{x\in R_x, y\in R_y} y f_{X,Y}(x,y) \\ \text{Why?} \quad &= \sum_{x\in R_x} x f_X(x) + \sum_{y\in R_y} y f_Y(y) \\ &= EX + EY. \end{aligned}$$

More generally, if $X_1, \ldots, X_N$ are random variables:

$E\left(\sum_{n=1}^{N} X_n\right) = \sum_{n=1}^{N} EX_n.$

To know the mean of the sum we don't need to know the joint distribution!

Consequences of law of unconscious statistician ii

If X, Y are independent: $E(X \cdot Y) = E(X) \cdot E(Y)$

$$\begin{aligned} E(X \cdot Y) &= \sum_{x \in R_x, y \in R_y} x \cdot y \cdot f_{X,Y}(x, y) \\ \text{Independence:} \quad &= \sum_{x \in R_x, y \in R_y} x \cdot y \cdot f_X(x) f_Y(y) \\ \text{Why?} \quad &= \sum_{x \in R_x} x f_X(x) \sum_{y \in R_y} y f_Y(y) \\ &= EX \cdot EY. \end{aligned}$$

Continuous Random Variables

This entire discussion on multiple random variables holds for continuous random varaibles by replacing sums with integrals.

- $\int_x \int_y f(x, y)dxdy = 1$
- $f_X(x) = \int_y f(x, y)dy$ is the marginal density of X
- $f_Y(y) = \int_x f(x, y)dx$ is the marginal density of Y
- $f_{Y|X}(y|X = x) = f(x, y)/f_X(x)$ is the conditional density of Y given $X = x$
- $f_{X|Y}(x|Y = y) = f(x, y)/f_Y(y)$ is the conditional density of X given $Y = y$
- $E[g(X, Y)] = \int_x \int_y g(x, y)f(x, y)dxdy$
- $E[X + Y] = E[X] + E[Y]$
- $E[XY] = E[X]E[Y]$ if X and Y are independent

Variance of a Random Variable

Variance of a random variable X with mean μ_X take $g(x) = (x - \mu_X)^2$.

$Var(X) = Eg(X)$.

Show that (on the chalk board):

$Var(X) = EX^2 - \mu_X^2$

$E(aX + b) = aE(X) + b$.

$Var(aX + b) = a^2 Var(X)$.

Covariance of two random variables

Covariance of two random variables X, Y with means μ_X, μ_Y.

Take $g(x, y) = (x - \mu_X) \cdot (y - \mu_Y)$.

$Cov(X, Y) = Eg(X, Y)$ - A measure of how the variables 'covary'.
$Cov(X, Y) > 0 \longrightarrow$ when X increases Y tends to increase.

$Cov(X, Y) < 0 \longrightarrow$ when X increases Y tends to decrease. Show

that $Cov(aX, bY) = a \cdot bCov(X, Y)$.

Changing the units of a measurement will change covariance.

Correlation of two random variables

Correlation $\rho(X, Y)$ does not depend on units of measurement:

$\rho(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

Show that: $\rho(aX, bY) = \rho(X, Y)$.

Independence

Independence:

Show that $Cov(X, Y) = EXY - \mu_X \mu_Y$.

Conclude: If X, Y are independent $Cov(X, Y) = 0$. (Converse is not true.)

Variance of a sum of random variables

$$
\begin{aligned}
Var(X+Y) =& E[(X+Y)^2] - [E(X+Y)]^2 \\
=& E[X^2 + 2XY + Y^2] - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2 \\
=& E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 \\
&+ 2E(XY) - 2E(X)E(Y) \\
=& Var(X) + Var(Y) + 2Cov(X,Y)
\end{aligned}
$$

Conclude: If X, Y independent: $Var(X+Y) = Var(X) + Var(Y)$.

More generally, if $X_1, \ldots, X_N$ are independent then

$Var(\sum_{n=1}^{N} X_n) = \sum_{n=1}^{N} Var(X_n)$.

Properties of sample average: Set up

We draw with replacement from a box containing tickets of numbers N times and record the number on each draw as $X_1, X_2, \ldots, X_N$. Because we draw with replacement we can *assume* that the variables X_n are independent.

Let μ_B be the *average of the box* and let σ_B^2 be the mean square deviation (MSD) of the box.

Recall that $EX_n = \mu_B$ and $VarX_n = \sigma_B^2$ for each n.

Denote the sample average as $\overline{X} = \frac{1}{N} \sum_{n=1}^{N} X_n$.

Mean of Sample Average

$$\begin{aligned} E\overline{X} &= \frac{1}{N} E \sum_{n=1}^{N} X_n \\ \text{Why?} \quad &= \frac{1}{N} \sum_{n=1}^{N} EX_n \\ \text{Why?} \quad &= \mu_B. \end{aligned}$$

Variance of Sample Average

The variance of the sample average

$$
\begin{aligned}
Var(\overline{X}) &= Var(\frac{1}{N}\sum_{n=1}^{N} X_n) \\
&= \frac{1}{N^2} Var(\sum_{n=1}^{N} X_n) \\
\text{Independence:} \quad &= \frac{1}{N^2}\sum_{n=1}^{N} Var(X_n) \\
&= \frac{\sigma_N^2}{N}
\end{aligned}
$$

Conclusion

So the expected value of the average of the sample is the average of the box no matter how large the sample.

The variance of the sample average *decreases* as the sample size N increases.

Set up: Proportions

Example: $X_1, \ldots, X_N$ are draws from a box of 0's and 1's with replacement. Assume fraction of 1's in the box is p.

Each X_n is Ber(p), i.e. $EX_n = p, Var(X_n) = p(1-p)$.

Since the draws are with replacement we can assume they are independent and so writing $S = \sum_{n=1}^{N} X_n$ we have that S is Binomial(N,p).

$f_S(n) = \binom{N}{n} p^n (1-p)^{N-n}$.

We can compute ES using the definition $\sum_{n=1}^{N} n \binom{N}{n} p^n (1-p)^{N-n}$ but that requires some complicated algebra.

Mean and variance of Sample Proportion

Instead we use the rules for mean and variance of a sum:

$E(S) = \sum_{n=1}^{N} EX_n = Np.$

And since X_n are independent:

$Var(S) = \sum_{n=1}^{N} Var(X_n) = Np(1-p).$

And for the sample average:

$E\overline{X} = E\frac{S}{N} = \frac{1}{N}ES = p.$

$Var(\overline{X}) = Var(\frac{S}{N}) = \frac{1}{N^2}Np(1-p) = \frac{p(1-p)}{N}.$

The expected value of the sample average is p - the proportion of 1's in the box: it is centered in the 'right' place.

The variance of the sample average decreases with sample size.