3.8 Functions of random variables
3.7, 3.9, 3.11 Multiple random variables (discrete)

Prof. Tesler

Math 186
Winter 2019

3.8 One random variable as a function of another random variable (Discrete)

- Let $X =$ roll of a biased die, $Y = 10X + 2$, and $Z = (X-3)^2$.

x	$p_X(x)$	$y = 10x+2$	$p_Y(y)$	$z = (x-3)^2$	$p_Z(z)$
1	q_1	12	q_1	4	
2	q_2	22	q_2	1	
3	q_3	32	q_3	0	$p_Z(0) = q_3$
4	q_4	42	q_4	1	$p_Z(1) = q_2 + q_4$
5	q_5	52	q_5	4	$p_Z(4) = q_1 + q_5$
6	q_6	62	q_6	9	$p_Z(9) = q_6$

- For $W = g(X)$, the pdf $p_W(w)$ is the sum of $p_X(x)$ over all possible inverses $x = g^{-1}(w)$, or $p_W(w) = 0$ if there are no inverses:

$$p_W(w) = \sum_{x\,:\,g(x)=w} p_X(x)$$

Function of a Discrete Random Variable

- Let $X =$ roll of a biased die, $Y = 10X + 2$, and $Z = (X-3)^2$.

x	$p_X(x)$	$y = 10x+2$	$p_Y(y)$	$z = (x-3)^2$	$p_Z(z)$
1	q_1	12	q_1	4	
2	q_2	22	q_2	1	
3	q_3	32	q_3	0	$p_Z(0) = q_3$
4	q_4	42	q_4	1	$p_Z(1) = q_2 + q_4$
5	q_5	52	q_5	4	$p_Z(4) = q_1 + q_5$
6	q_6	62	q_6	9	$p_Z(9) = q_6$

- For $Y = aX + b$ with $a \neq 0$, the unique inverse is $X = \frac{Y-b}{a}$, so

$$p_Y(y) = p_X\left(\frac{y-b}{a}\right)$$

Function of a Discrete Random Variable

- Let $Y = X^2$.
- Inverses of $y = x^2$ are

$$x = \begin{cases} \pm\sqrt{y} & \text{if } y > 0; \\ 0 & \text{if } y = 0; \\ \text{none} & \text{if } y < 0. \end{cases}$$

- $$P(Y = y) = P(X^2 = y) = \begin{cases} P(X = \sqrt{y}) + P(X = -\sqrt{y}) & \text{if } y > 0; \\ P(X = 0) & \text{if } y = 0; \\ 0 & \text{if } y < 0. \end{cases}$$

- Use pdf notation to express pdf of Y in terms of pdf of X:

$$p_Y(y) = \begin{cases} p_X(\sqrt{y}) + p_X(-\sqrt{y}) & \text{if } y > 0; \\ p_X(0) & \text{if } y = 0; \\ 0 & \text{if } y < 0. \end{cases}$$

Function of a Continuous Random Variable

- Let U be uniform on the real interval $[1, 7]$, so

$$f_U(u) = \begin{cases} 1/6 & \text{if } 1 \leqslant u \leqslant 7 \text{ (in real numbers);} \\ 0 & \text{otherwise} \end{cases}$$

- Let $V = 2U$. This is uniform on $[2, 14]$, so

$$f_V(v) = \begin{cases} 1/12 & \text{if } 2 \leqslant v \leqslant 14 \text{ (in real numbers);} \\ 0 & \text{otherwise} \end{cases}$$

- Even though there's just one inverse for each value, the probabilities of corresponding values are different!

Continuous Random Variables: Computing $f_Y(y)$ from $f_X(x)$ when $Y = g(X)$

Procedure

- Let $Y = g(X)$.
- Compute $F_X(x)$.
- Compute $F_Y(y) = P(Y \leqslant y) = P(g(X) \leqslant y) = \cdots$
  - Details depend on the function.
  - Typically it is expressed in terms of $F_X(x)$ at various values of x (the ones where $g(x) = y$).
- Compute $f_Y(y) = \dfrac{d}{dy} F_Y(y)$.

Example with one-to-one function

- Let $f_X(x) = 8e^{-8x}$ for $x \geqslant 0$.
- It's a valid pdf since it's $\geqslant 0$ and the total probability is 1:

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^{\infty} 8e^{-8x}\,dx = \left.\frac{8e^{-8x}}{-8}\right|_{x=0}^{\infty}$$
$$= -(e^{-\infty} - e^0) = -(0-1) = 1$$

- The CDF for $x \geqslant 0$ is

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_0^{x} 8e^{-8t}\,dt = \left.\frac{8e^{-8t}}{-8}\right|_{t=0}^{x}$$
$$= -(e^{-8x} - e^0) = 1 - e^{-8x}$$

while for $x < 0$, the CDF is $F_X(x) = 0$.

Example with one-to-one function

pdf: $f_X(x) = \begin{cases} 8e^{-8x} & \text{if } x \geqslant 0; \\ 0 & \text{if } x < 0 \end{cases}$ **cdf:** $F_X(x) = \begin{cases} 1 - e^{-8x} & \text{if } x \geqslant 0; \\ 0 & \text{if } x < 0 \end{cases}$

- We'll compute the pdf of $Y = 10X + 2$.
- First, we convert the two cases $x \geqslant 0$ and $x < 0$ to y.
- Note $x \geqslant 0$ gives $10x \geqslant 0$,
  so $10x + 2 \geqslant 2$,
  so $y \geqslant 2$.
- Similarly, $x < 0$ gives $y < 2$.

Example with one-to-one function

$$f_X(x) = \begin{cases} 8e^{-8x} & \text{if } x \geqslant 0; \\ 0 & \text{if } x < 0 \end{cases} \qquad F_X(x) = \begin{cases} 1 - e^{-8x} & \text{if } x \geqslant 0; \\ 0 & \text{if } x < 0 \end{cases} \qquad Y = 10X + 2$$

Wrong Way: $f_Y(y) = f_X(g^{-1}(y))$ as in the discrete case

- Try $f_Y(y) = f_X((y-2)/10)$:

$$f_Y(y) = \begin{cases} 8e^{-8(y-2)/10} & \text{if } y \geqslant 2; \\ 0 & \text{if } y < 2. \end{cases}$$

- Compute the total probability:

$$\int_{-\infty}^{\infty} f_Y(y)\,dy = \int_2^{\infty} 8e^{-8(y-2)/10}\,dy = \left.\frac{8e^{-8(y-2)/10}}{-8/10}\right|_{y=2}^{\infty}$$

$$= -10(e^{-\infty} - e^0) = -10(0-1)$$

$$= 10 \neq 1$$

- The total probability is not 1, so this is not a valid pdf.

Example with one-to-one function

$$f_X(x) = \begin{cases} 8e^{-8x} & \text{if } x \geqslant 0; \\ 0 & \text{if } x < 0 \end{cases} \qquad F_X(x) = \begin{cases} 1 - e^{-8x} & \text{if } x \geqslant 0; \\ 0 & \text{if } x < 0 \end{cases} \qquad Y = 10X + 2$$

Right way: Compute CDF $F_Y(y)$ and differentiate $f_Y(y) = F_Y{}'(y)$

- Compute CDF $F_Y(y)$:
  $F_Y(y) = P(Y \leqslant y) = P(10X + 2 \leqslant y)$
  $$= P\left(X \leqslant \frac{y-2}{10}\right) = F_X\left(\frac{y-2}{10}\right) = \begin{cases} 1 - e^{-8(y-2)/10} & \text{if } y \geqslant 2; \\ 0 & \text{if } y < 2. \end{cases}$$
- Differentiate to get PDF $f_Y(y)$:
  $$f_Y(y) = F_Y{}'(y) = \begin{cases} \frac{8}{10} e^{-8(y-2)/10} & \text{if } y > 2; \\ 0 & \text{if } y < 2. \end{cases}$$

$Y = aX + b$ in general

where a, b are constants and $a \neq 0$

- In general, for a continuous random variable X, the pdfs of X and $Y = aX + b$ are related by

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

- Whereas if X is a discrete random variable, then

$$p_Y(y) = p_X\left(\frac{y-b}{a}\right)$$

Note the scaling factor $1/|a|$ for the continuous case, but not for the discrete case.

Example with general function $Y = g(X)$

- Compute CDF and PDF of $Y = X^2$ where X is a continuous R.V.
- When $y \leqslant 0$,
  - $F_Y(y) = P(Y \leqslant y) = P(X^2 \leqslant y) = 0$
  - $f_Y(y) = \dfrac{d}{dy} F_Y(y) = \dfrac{d}{dy} 0 = 0$
- When $y > 0$,
  - $$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(X^2 \leqslant y) \\ &= P(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$
  - $$\begin{aligned} f_Y(y) &= \frac{d}{dy}\left(F_X(\sqrt{y}) - F_X(-\sqrt{y})\right) \\ &= {F_X}'(\sqrt{y})\left(\tfrac{1}{2\sqrt{y}}\right) - {F_X}'(-\sqrt{y})\left(-\tfrac{1}{2\sqrt{y}}\right) \\ &= \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} \end{aligned}$$

- Consider a biased 6-sided die where q_i is the probability of rolling i for $i = 1, 2, \ldots, 6$. (6 sides is an example, it could be any # sides.) Each q_i is between 0 and 1, and $q_1 + \cdots + q_6 = 1$.
- The probability of a sequence of independent rolls is

$$P(1131326) = q_1\,q_1\,q_3\,q_1\,q_3\,q_2\,q_6 = {q_1}^3\,q_2\,{q_3}^2\,q_6 = \prod_{i=1}^{6} {q_i}^{\#\ i\text{'s}}$$

- Roll the die n times ($n = 0, 1, 2, 3, \ldots$).
  Let X_1 be the number of 1's, X_2 be the number of 2's, etc.

$$p_{X_1,X_2,\ldots,X_6}(k_1,k_2,\ldots,k_6) = P(X_1 = k_1, X_2 = k_2, \ldots, X_6 = k_6)$$

$$= \begin{cases} \binom{n}{k_1,k_2,\ldots,k_6} {q_1}^{k_1} {q_2}^{k_2} \ldots {q_6}^{k_6} & \text{if } k_1,\ldots,k_6 \text{ are integers} \geqslant 0 \text{ adding up to } n; \\ 0 & \text{otherwise.} \end{cases}$$

Genetics example

Consider a $TtRR \times TtRr$ cross of pea plants:

Punnett Square

	TR $(1/2)$	tR $(1/2)$
TR $(1/4)$	$TTRR$ $(1/8)$	$TtRR$ $(1/8)$
Tr $(1/4)$	$TTRr$ $(1/8)$	$TtRr$ $(1/8)$
tR $(1/4)$	$TtRR$ $(1/8)$	$ttRR$ $(1/8)$
tr $(1/4)$	$TtRr$ $(1/8)$	$ttRr$ $(1/8)$

Genotype	**Prob.**
$TTRR$	$1/8$
$TtRR$	$2/8 = 1/4$
$TTRr$	$1/8$
$TtRr$	$2/8 = 1/4$
$ttRR$	$1/8$
$ttRr$	$1/8$

Genetics example

- If there are 27 offspring, what is the probability that 9 offspring have genotype TTRR, 2 have genotype TtRR, 3 have genotype TTRr, 5 have genotype TtRr, 7 have genotype ttRR, and 1 has genotype ttRr?
- Use the multinomial distribution:

Genotype	Probability	Frequency
TTRR	1/8	9
TtRR	1/4	2
TTRr	1/8	3
TtRr	1/4	5
ttRR	1/8	7
ttRr	1/8	1
Total	1	27

$$P = \frac{27!}{9!\,2!\,3!\,5!\,7!\,1!}\left(\frac{1}{8}\right)^9\left(\frac{1}{4}\right)^2\left(\frac{1}{8}\right)^3\left(\frac{1}{4}\right)^5\left(\frac{1}{8}\right)^7\left(\frac{1}{8}\right)^1 \approx 2.19 \cdot 10^{-7}$$

Genetics example

- If there are 25 offspring, what is the probability that 9 offspring have genotype TTRR, 2 have genotype TtRR, 3 have genotype TTRr, 5 have genotype TtRr, 7 have genotype ttRR, and 1 has genotype ttRr?
- $P = 0$ because the numbers $9, 2, 3, 5, 7, 1$ do not add up to 25.

Joint distribution — Example (Discrete Case)

- Roll a fair 6-sided die n times.
- Let $X =$ # of 1's. X has a binomial distribution with probability $\frac{1}{6}$:

$$p_X(x) = \begin{cases} \binom{n}{x}(1/6)^x(5/6)^{n-x} & \text{if } x = 0, 1, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}$$

- Let $Y =$ # of 2's and 3's. Y has a binomial distribution with probability $\frac{2}{6} = \frac{1}{3}$:

$$p_Y(y) = \begin{cases} \binom{n}{y}(1/3)^y(2/3)^{n-y} & \text{if } y = 0, 1, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}$$

- There are x 1's, y 2's and 3's, and $n - x - y$ other numbers (4,5,6), so the joint pdf of X and Y is multinomial:

$$p_{X,Y}(x,y) = \begin{cases} \binom{n}{x,y,n-x-y}(1/6)^x(2/6)^y(3/6)^{n-x-y} \\ \qquad \text{provided } x, y \text{ are integers} \geqslant 0 \text{ and } x + y \leqslant n; \\ 0 \quad \text{otherwise.} \end{cases}$$

Joint distribution — Probability Density Function

We will work with the joint pdf of X and Y for 3 rolls:

$$p_{X,Y}(x,y) = \begin{cases} \binom{3}{x,y,3-x-y}(1/6)^x(2/6)^y(3/6)^{3-x-y} \\ \quad \text{provided } x,y \text{ are integers } \geqslant 0 \text{ and } x+y \leqslant 3; \\ 0 \quad \text{otherwise.} \end{cases}$$

$p_{X,Y}(x,y)$	$x=0$	1	2	3
$y=0$	$1/8$	$1/8$	$1/24$	$1/216$
1	$1/4$	$1/6$	$1/36$	0
2	$1/6$	$1/18$	0	0
3	$1/27$	0	0	0

Joint distribution — Marginal densities

-

$p_{X,Y}(x,y)$	$x=0$	1	2	3	Total $p_Y(y)$
$y=0$	$1/8$	$1/8$	$1/24$	$1/216$	$8/27$
1	$1/4$	$1/6$	$1/36$	0	$4/9$
2	$1/6$	$1/18$	0	0	$2/9$
3	$1/27$	0	0	0	$1/27$
Total $p_X(x)$	$125/216$	$25/72$	$5/72$	$1/216$	1

- **Total of column x:** $p_X(x) = \sum_y p_{X,Y}(x,y)$
- **Total of row y:** $p_Y(y) = \sum_x p_{X,Y}(x,y)$
- $p_X(x)$ and $p_Y(y)$ are the marginal densities of $p_{X,Y}(x,y)$ because they are in the margins of the table.
- **Grand total of table:** $\sum_x \sum_y p_{X,Y}(x,y) = 1$

Joint distribution — Marginal densities

- With multiple variables, the marginal density for each variable is obtained by summing over the others:

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

$$p_Y(y) = \sum_x \sum_z p_{X,Y,Z}(x,y,z)$$

$$p_Z(z) = \sum_x \sum_y p_{X,Y,Z}(x,y,z)$$

- Summing over one or more variables gives the joint PDF in the remaining variable(s):

$$p_{X,Y}(x,y) = \sum_z p_{X,Y,Z}(x,y,z)$$

Joint distribution — Functions

- The pdf of $Z = g(X, Y)$ is
$$p_Z(z) = \sum_{(x,y)\,:\,g(x,y)=z} p_{X,Y}(x, y)$$
- For $Z = X + Y$, this is
$$p_Z(z) = \sum_x p_{X,Y}(x, z - x)$$
- $$p_Z(2) = p_{X,Y}(0, 2) + p_{X,Y}(1, 1) + p_{X,Y}(2, 0) = \frac{1}{6} + \frac{1}{6} + \frac{1}{24} = \frac{3}{8}$$

$p_{X,Y}(x, y)$	$x = 0$	1	2	3
$y = 0$	1/8	1/8	1/24	1/216
1	1/4	1/6	1/36	0
2	1/6	1/18	0	0
3	1/27	0	0	0

Joint distribution — Conditional densities

- Suppose you know that $Y = 2$.

$p_{X,Y}(x,y)$	$x = 0$	1	2	3	Total $p_Y(y)$
$y = 0$	$1/8$	$1/8$	$1/24$	$1/216$	$8/27$
1	$1/4$	$1/6$	$1/36$	0	$4/9$
2	$1/6$	$1/18$	0	0	$2/9$
3	$1/27$	0	0	0	$1/27$
Total $p_X(x)$	$125/216$	$25/72$	$5/72$	$1/216$	1

-

$$P(X = 0|Y = 2) = \frac{P(X=0,Y=2)}{P(Y=2)} = \frac{1/6}{2/9} = \frac{3}{4}$$

$$P(X = 1|Y = 2) = \frac{P(X=1,Y=2)}{P(Y=2)} = \frac{1/18}{2/9} = \frac{1}{4}$$

$$P(X = x|Y = 2) = \frac{P(X=x,Y=2)}{P(Y=2)} = \frac{0}{2/9} = 0 \qquad \text{if } x \neq 0, 1$$

Joint distribution — Conditional densities

- **Our book's notation:**

$$\begin{aligned} p_{X|Y=y}(x) = p_{X|y}(x) &= P(X = x|Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)} \end{aligned}$$

$$p_{X|2}(0) = 3/4 \qquad p_{X|2}(1) = 1/4 \qquad p_{X|2}(x) = 0 \text{ if } x \neq 0, 1$$

- Conditional density in column x instead of row y:

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

Joint distribution — Independence

Discrete random variables $X, Y, Z, \ldots$ are independent if

$$p_{X,Y,Z}(x, y, z) = p_X(x)\, p_Y(y)\, p_Z(z) \quad \text{for } \textit{all values} \text{ of } x, y, z.$$

If there are any exceptions, they are not independent.

Example (Repeated trials)

- A biased die has probability q_i of showing $i = 1, \ldots, 6$.
- U, V are the values of two independent rolls.
- $p_{U,V}(i,j) = \begin{cases} q_i\, q_j & \text{for } i \text{ and } j \text{ in } 1, \ldots, 6; \\ 0 & \text{otherwise.} \end{cases}$
- U and V are independent.
- For n independent rolls $U_1, \ldots, U_n$, the probability is multiplicative, so $U_1, \ldots, U_n$ are independent random variables.

Joint distribution — Independence

- In our running example, check $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ for all (x,y):

$p_{X,Y}(x,y)$	$x=0$	1	2	3	Total $p_Y(y)$
$y=0$	$1/8$	$1/8$	$1/24$	$1/216$	$8/27$
1	$1/4$	$1/6$	$1/36$	0	$4/9$
2	$1/6$	$1/18$	0	0	$2/9$
3	$1/27$	0	0	0	$1/27$
Total $p_X(x)$	$125/216$	$25/72$	$5/72$	$1/216$	1

- $p_X(3)p_Y(2) = (1/216)(2/9) = 1/1072$ but $p_{X,Y}(3,2) = 0$.
  (There are other exceptions too, but it only takes one.)
  So X and Y are *not* independent.

Joint Cumulative Distribution Function (CDF)

The joint cumulative distribution function (cdf) for X, Y, Z is

$$\begin{aligned} F_{X,Y,Z}(x,y,z) &= P(X \leqslant x, Y \leqslant y, Z \leqslant z) \\ &= \sum_{u \leqslant x} \sum_{v \leqslant y} \sum_{w \leqslant z} p_{X,Y,Z}(u,v,w) \end{aligned}$$

Joint Cumulative Distribution Function (CDF)

- $F_{X,Y}(1,2) = 8/9$ by summing $p_{X,Y}(x,y)$ for $x = 0, 1$, $y = 0, 1, 2$:

pdf $p_{X,Y}(x,y)$	$x = 0$	1	2	3
$y = 0$	$1/8$	$1/8$	$1/24$	$1/216$
1	$1/4$	$1/6$	$1/36$	0
2	$1/6$	$1/18$	0	0
3	$1/27$	0	0	0

cdf $F_{X,Y}(x,y)$	$x = 0$	1	2	3
$y = 0$	$1/8$	$1/4$	$7/24$	$8/27$
1	$3/8$	$2/3$	$53/72$	$20/27$
2	$13/24$	$8/9$	$23/24$	$26/27$
3	$125/216$	$25/27$	$215/216$	1

Joint Cumulative Distribution Function (CDF)

Values of X, Y not in the range

cdf $F_{X,Y}(x,y)$	$x=0$	1	2	3
$y=0$	$1/8$	$1/4$	$7/24$	$8/27$
1	$3/8$	$2/3$	$53/72$	$20/27$
2	$13/24$	$8/9$	$23/24$	$26/27$
3	$125/216$	$25/27$	$215/216$	1

- The pdf is 0 for values of (X,Y) not in the range:
  $$p_{X,Y}(1.5, 2.3) = 0$$
  But the cdf is usually NOT 0:
  $$F_{X,Y}(1.5, 2.3) = P(X \leqslant 1.5, Y \leqslant 2.3) = P(X \leqslant 1, Y \leqslant 2) = F_{X,Y}(1,2) = \frac{8}{9}$$
- $F_{X,Y}(-1, 2.5) = P(X \leqslant -1, Y \leqslant 2.5) = 0$
- $F_{X,Y}(10, 0) = P(X \leqslant 10, Y \leqslant 0) = P(X \leqslant 3, Y \leqslant 0) = F_{X,Y}(3,0) = 8/27$

Expected Values for Joint pdfs

- Let $Z = g(X, Y)$ be a function of discrete random variables X, Y. The expected value of $g(X, Y)$ is

$$E(g(X,Y)) = \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y)$$

- If there are more random variables, use more Σ's.
- $E(g(X, Y))$ has the same value as $E(Z)$:

$$E(Z) = \sum_z z\, p_Z(z)$$

- $E(g(X, Y))$ estimates the average

$$\frac{g(x_1, y_1) + \cdots + g(x_n, y_n)}{n}$$

where $(x_1, y_1), \ldots, (x_n, y_n)$ are randomly chosen with probabilities given by the joint distribution.

Expected Values for Joint pdfs

-

pdf $p_{X,Y}(x,y)$	$x=0$	1	2	3
$y=0$	$1/8$	$1/8$	$1/24$	$1/216$
1	$1/4$	$1/6$	$1/36$	0
2	$1/6$	$1/18$	0	0
3	$1/27$	0	0	0

- $$\begin{aligned} E(X+Y) = \\ &(0+0)(1/8) &+(1+0)(1/8) &+(2+0)(1/24) &+(3+0)(1/216) \\ +&(0+1)(1/4) &+(1+1)(1/6) &+(2+1)(1/36) &+(3+1)(0) \\ +&(0+2)(1/6) &+(1+2)(1/18) &+(2+2)(0) &+(3+2)(0) \\ +&(0+3)(1/27) &+(1+3)(0) &+(2+3)(0) &+(3+3)(0) \\ &= 3/2 \end{aligned}$$

- Next we'll show an easier way to compute this:

$$\begin{aligned} E(X+Y) &= E(X)+E(Y) \\ &= 3(1/6)+3(2/6) = (1/2)+1 = 3/2 \end{aligned}$$

(recall that X and Y separately have binomial distributions)

Expected Values for Joint pdfs — Properties

Theorem

$E(X+Y) = E(X) + E(Y)$, provided $E(X)$ and $E(Y)$ are defined and finite.

Proof (discrete case).

$$E(X+Y) = \sum_x \sum_y (x+y) p_{X,Y}(x,y)$$

$$= \sum_x x \left(\sum_y p_{X,Y}(x,y) \right) + \sum_y y \left(\sum_x p_{X,Y}(x,y) \right)$$

(inside sums evaluate to marginal densities)

$$= \sum_x x\, p_X(x) + \sum_y y\, p_Y(y)$$

$$= E(X) + E(Y) \qquad \square$$

Expected Values for Joint pdfs — Properties

These addition properties hold for both discrete and continuous random variables:

- $E(X+Y+Z+\cdots) = E(X)+E(Y)+E(Z)+\cdots$
- $E(aX+bY+cZ+\cdots) = a\,E(X)+b\,E(Y)+c\,E(Z)+\cdots$ for constants $a, b, c, \ldots$.
- $E\left(g_1(X,Y,Z)+g_2(X,Y,Z)+\cdots\right)$
  $= E\left(g_1(X,Y,Z)\right)+E\left(g_2(X,Y,Z)\right)+\cdots$

If X, Y, Z are independent, **there are two more properties:**

- $E(XYZ) = E(X)E(Y)E(Z)$
- $\mathrm{Var}(X+Y+Z) = \mathrm{Var}(X)+\mathrm{Var}(Y)+\mathrm{Var}(Z)$
- Those formulas extend to any number of independent variables.

If they are not independent, those properties may or may not hold.

Expected Value of a Product — Dependent Variables

Example (Dependent)

- Let U be the roll of a fair 6-sided die.
- Let V be the value of the exact same roll of the die ($U = V$).
- $E(U) = E(V) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2}$ and $E(U)E(V) = \frac{49}{4}$.
- $$E(UV) = \frac{1\cdot1+2\cdot2+3\cdot3+4\cdot4+5\cdot5+6\cdot6}{6} = \frac{91}{6}$$

Example (Independent)

- Now let U, V be the values of two independent rolls of a fair 6-sided die.
- $$E(UV) = \sum_{x=1}^{6}\sum_{y=1}^{6} \frac{x \cdot y}{36} = \frac{441}{36} = \frac{49}{4}$$

  and $E(U)E(V) = (7/2)(7/2) = 49/4$

Expected Value of a Product — Independent Variables

Proof of $E(XY) = E(X)E(Y)$ provided X, Y are independent:

$$
\begin{aligned}
E(XY) &= \sum_x \sum_y x\,y\,p_{X,Y}(x,y) \\
&= \sum_x \sum_y x\,y\,p_X(x)p_Y(y) \qquad \text{provided } X, Y \text{ are independent} \\
&= \left(\sum_x x\,p_X(x)\right)\left(\sum_y y\,p_Y(y)\right) \\
&= E(X)E(Y) \quad \square
\end{aligned}
$$

Variance of a Sum — Dependent Variables

- We will show that if X, Y are independent, then

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$$

Example (Dependent)

First consider this dependent example:
Let X be any non-constant random variable and $Y = -X$.

$$\begin{aligned}
\operatorname{Var}(X+Y) &= \operatorname{Var}(0) = 0 \\
\operatorname{Var}(X) + \operatorname{Var}(Y) &= \operatorname{Var}(X) + \operatorname{Var}(-X) \\
&= \operatorname{Var}(X) + (-1)^2 \operatorname{Var}(X) = 2\operatorname{Var}(X)
\end{aligned}$$

but usually $\operatorname{Var}(X) \neq 0$ (the only exception would be if X is a constant).

Variance of a Sum — Independent Variables

Theorem

If X, Y are independent, then $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Proof.

For any random variables X and Y:

$$E((X+Y)^2) = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2)$$
$$(E(X+Y))^2 = (E(X)+E(Y))^2 = (E(X))^2 + 2E(X)E(Y) + (E(Y))^2$$

$$\begin{aligned}
\mathrm{Var}(X+Y) &= E((X+Y)^2) - (E(X+Y))^2 \\
&= \left(E(X^2) - (E(X))^2\right) \\
&\qquad + 2\left(E(XY) - E(X)E(Y)\right) \\
&\qquad + \left(E(Y^2) - (E(Y))^2\right) \\
&= \mathrm{Var}(X) + 2(E(XY) - E(X)E(Y)) + \mathrm{Var}(Y)
\end{aligned}$$

If X, Y are independent then $E(XY) = E(X)E(Y)$, so the middle term vanishes. □

Variance of a Sum — Dependent Variables

- We showed that for any random variables X, Y (possibly dependent),

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2(E(XY) - E(X)E(Y))$$

- The *covariance* of X and Y is $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$.
  Rewrite the above as:

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y).$$

- If X and Y are independent, $\mathrm{Cov}(X,Y) = 0$.
  But $\mathrm{Cov}(X,Y) = 0$ does not guarantee independence.
  If $\mathrm{Cov}(X,Y) \neq 0$, then X, Y are dependent.

Mean and Variance of the Binomial Distribution

- A Bernoulli trial is a single flip of a coin, heads with probability p.
- Do n coin flips (n Bernoulli trials).

  Let X_i be the number of heads on flip i: $X_i = \begin{cases} 1 & \text{flip } i \text{ is heads;} \\ 0 & \text{flip } i \text{ is tails.} \end{cases}$
- The total number of heads in all flips is $X = X_1 + X_2 + \cdots + X_n$.
- The variables $X_1, \ldots, X_n$ are independent and have the same pdfs. They are called i.i.d. random variables, for independent identically distributed.
- $$\begin{aligned} E(X_1) &= 0(1-p) + 1p = p \\ E({X_1}^2) &= 0^2(1-p) + 1^2 p = p \\ \operatorname{Var}(X_1) &= E({X_1}^2) - (E(X_1))^2 = p - p^2 = p(1-p) \end{aligned}$$
- $E(X_i) = p$ and $\operatorname{Var}(X_i) = p(1-p)$ for all $i = 1, \ldots, n$ because they are identically distributed.

Mean and Variance of the Binomial Distribution

- The total number of heads in all flips is $X = X_1 + X_2 + \cdots + X_n$.
- $E(X_i) = p$ and $\text{Var}(X_i) = p(1-p)$ for all $i = 1, \ldots, n$.

Mean:

$$\begin{aligned} \mu_X = E(X) &= E(X_1 + \cdots + X_n) \\ &= E(X_1) + \cdots + E(X_n) \\ &= p + \cdots + p = np \quad \text{identically distributed} \end{aligned}$$

Variance:

$$\begin{aligned} {\sigma_X}^2 = \text{Var}(X) &= \text{Var}(X_1 + \cdots + X_n) \\ &= \text{Var}(X_1) + \cdots + \text{Var}(X_n) \quad \text{by independence} \\ &= p(1-p) + \cdots + p(1-p) \quad \text{identically distributed} \\ &= np(1-p) = npq \end{aligned}$$

Standard deviation:

$$\sigma_X = \sqrt{np(1-p)} = \sqrt{npq}$$