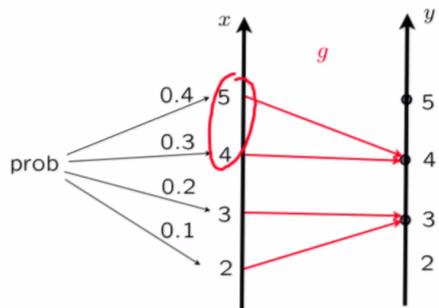


Unit 6 Further Topics on Random Variables

PMF of a function

Derived distributions — the discrete case

$$Y = g(X)$$

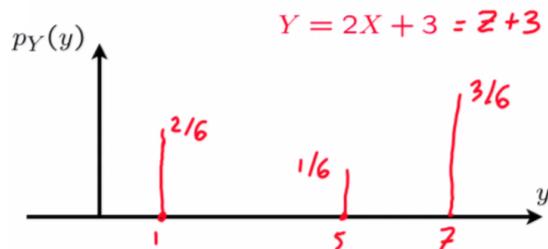
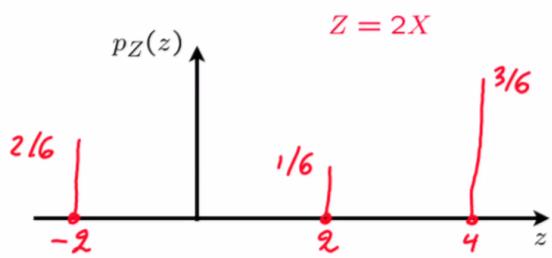
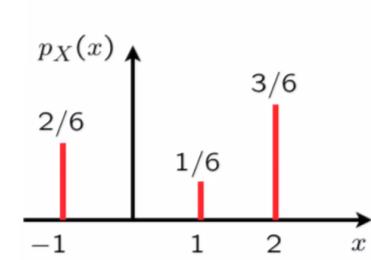


$$\begin{aligned} p_Y(4) &= P(Y=4) \\ &= P(X=4) + P(X=5) \\ &= p_X(4) + p_X(5) = 0.3 + 0.4 \end{aligned}$$

$$\begin{aligned} p_Y(y) &= P(g(X)=y) \\ &= \sum_{x: g(x)=y} p_X(x) \end{aligned}$$

Only sum the x's that give rise to the value of y that we are interested in

A linear function of a discrete r.v.



$$\begin{aligned} p_Y(y) &= P(Y=y) = P(2X+3=y) \\ &= P\left(X=\frac{y-3}{2}\right) = p_X\left(\frac{y-3}{2}\right) \end{aligned}$$

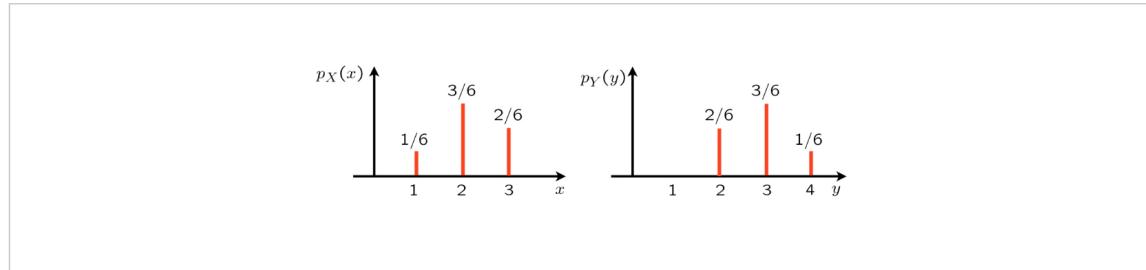
$$Y = aX + b : \quad p_Y(y) = p_X\left(\frac{y-b}{a}\right)$$

To find the PMF of a linear function we 'stretch' it by the multiple of x and shift it by the added constant

Exercise: Linear functions of discrete r.v.'s

2/2 points (graded)

The random variables X and Y obey a linear relation of the form $Y = aX + b$ and have the PMFs shown in the diagram. Find the values of a and b .



$$a = \boxed{-1}$$

✓ Answer: -1

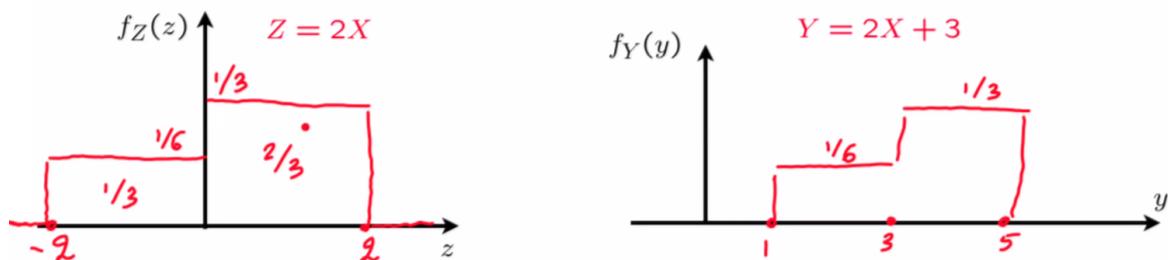
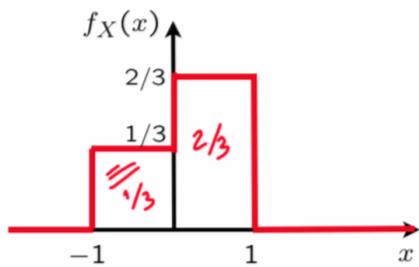
$$b = \boxed{5}$$

✓ Answer: 5

Solution:

Because the entries of the PMF of Y appear in the opposite order than the entries of the PMF of X , we know that a has to be negative. Furthermore, the spread of the PMF of Y is the same as the spread of the PMF of X , and therefore, $a = -1$. The random variable $-X$ takes values in the set $\{-3, -2, -1\}$. To obtain the given PMF of Y , we need to shift it (to the right) by $b = 5$.

A linear function of a continuous r.v.



When multiplying the area is stretched out but needs to remain the same total value, therefore the height of the bars lowers by a factor of the multiplier of X (2 here)

A linear function of a continuous r.v.

$$Y = aX + b$$

$a > 0$

$$\mathbb{P}(Y=y) = \mathbb{P}(aX+b=y) = \mathbb{P}\left(X=\frac{y-b}{a}\right)$$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(aX+b \leq y)$$

$$= \mathbb{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

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

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

$a < 0$

$$= \mathbb{P}\left(X \geq \frac{y-b}{a}\right)$$

$$= 1 - \mathbb{P}\left(X \leq \frac{y-b}{a}\right)$$

$$= 1 - F_X\left(\frac{y-b}{a}\right)$$

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

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

Work with CDF (red) then differentiate to get the PDF

$1/|a|$ ensures the formula integrates to 1

Continuous(left) discrete(right) without scaling

A linear function of a normal r.v. is normal

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

$$Y = aX + b \quad a \neq 0$$

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|} \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{y-b}{a}-\mu\right)^2/2\sigma^2} \\ &= \frac{1}{\sqrt{2\pi}\sigma|a|} e^{-\frac{(y-b-a\mu)^2}{2\sigma^2 a^2}} \end{aligned}$$

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

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Start with a bell-shaped normal PDF (centred at 0) then scale and shift it but still have a bell-shaped PDF

(a) Let X be an exponential random variable and let $Y = aX + b$. The random variable Y is exponential if and only if (choose one of the following statements):

always.

$a \neq 0$.

$a \neq 0$ and $b = 0$

$a > 0$

$a > 0$ and $b = 0$ ✓

$a = 1$

✗

(b) Let X be a continuous random variable, uniformly distributed on some interval, and let $Y = aX + b$. The random variable Y will be a continuous random variable with a uniform distribution if and only if (choose one of the following statements):

always.

$a > 0$.

$a \neq 0$

$a \neq 0$ and $b = 0$

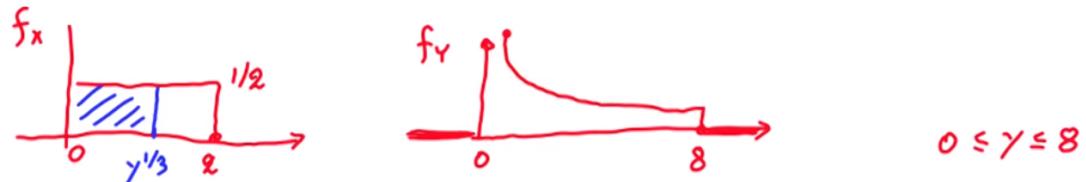
✓

Solution:

(a) For Y to be exponential, its range must be $[0, \infty)$. This will be the case only if $a > 0$ and $b = 0$. And if indeed $a > 0$ and $b = 0$, and X has parameter λ , then, for $y \geq 0$, $f_Y(y) = (1/a) f_X(y/a) = (\lambda/a) e^{-\lambda y/a}$, which is exponential (with parameter λ/a).

(b) A scaled and shifted uniform is uniform, except that if $a = 0$, then Y is a constant random variable, and therefore no longer continuous.

Example: $Y = X^3$; X uniform on $[0, 2]$



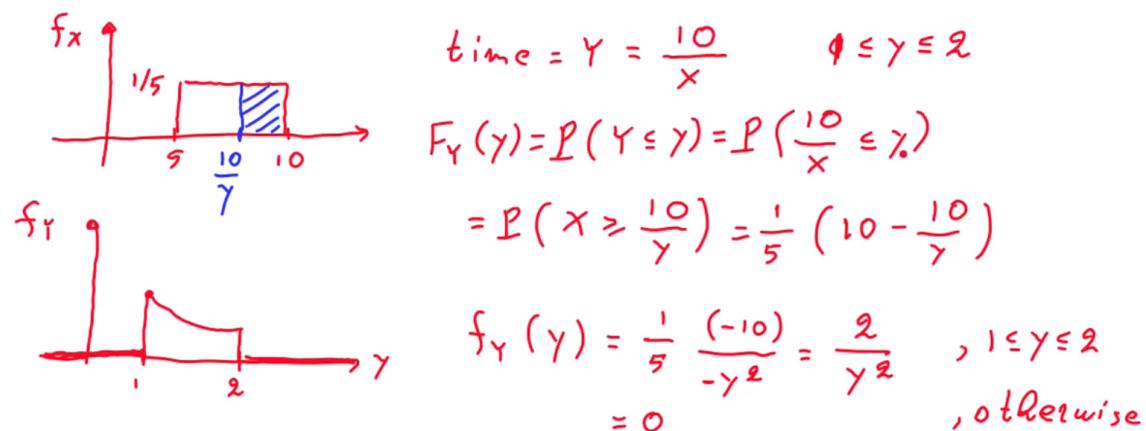
$$F_Y(y) = P(Y \leq y) = P(X^3 \leq y) = P(X \leq y^{1/3}) = \frac{1}{2} y^{1/3}$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{2} \cdot \frac{1}{3} y^{-2/3} = \frac{1}{6} \cdot \frac{1}{y^{2/3}}$$

second step is differentiation of the CDF in the first step to get the PDF

Example: $Y = a/X$

- You go to the gym and set the speed X of the treadmill to a number between 5 and 10 km/hr (with a uniform distribution). Find the PDF of the time it takes to run 10km.



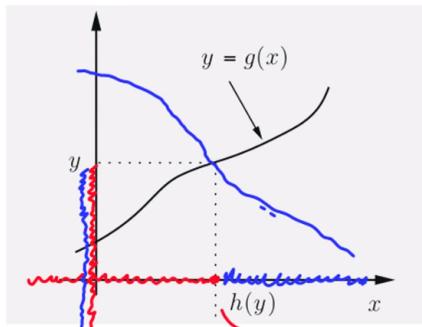
$\frac{1}{5}$ is height of CDF because $10 - 5 = 5$ and $5 * \frac{1}{5} = 1$

We differentiate the CDF to get the PDF

A general formula for the PDF of $Y = g(X)$ when g is monotonic $x^3 \frac{\alpha}{x}$
~~decreasing~~ $x < x' \Rightarrow g(x) < g(x')$

Assume g strictly increasing

and differentiable



inverse function $h \rightarrow$ decreasing

$$F_Y(y) = P(Y \leq y) = P(X \leq h(y)) = F_X(h(y))$$

$$f_Y(y) = f_X(h(y)) \left| \frac{d h}{d y}(y) \right|$$

$$F_Y(y) = P(Y \leq y) = P(X \geq h(y)) \\ = 1 - P(X \leq h(y)) = 1 - F_X(h(y))$$

$$f_Y(y) = f_X(h(y)) \left| \frac{d h}{d y}(y) \right|$$

$$f_Y(y) = f_X(h(y)) \left| \frac{d h}{d y}(y) \right|$$

monotonic - constantly increasing/decreasing

$g(x)$ takes us from x to y

$h(y)$ takes us from y to x and is known as the inverse function

Example: $Y = X^2$; X uniform on $[0, 1]$

$$f_Y(y) = f_X(h(y)) \left| \frac{d h}{d y}(y) \right|$$



$$y = x^2 \Leftrightarrow x = \sqrt{y} \quad h(y) = \sqrt{y}$$

$$f_Y(y) = \frac{1}{2\sqrt{y}}$$

$$0 \leq y \leq 1$$

The random variable X has a PDF of the form

$$f_X(x) = \begin{cases} \frac{1}{x^2}, & \text{for } x \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Let $Y = X^2$. For $y \geq 1$, the PDF of Y it takes the form $f_Y(y) = \frac{a}{y^b}$. Find the values of a and b .

$a =$ ✖ Answer: 0.5

$b =$ ✖ Answer: 1.5

Solution:

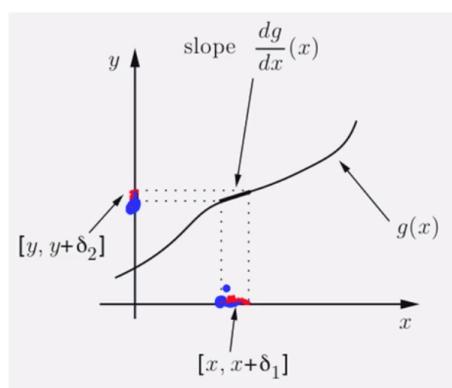
For any $y \geq 1$, we have

$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(X^2 \leq y) = \mathbf{P}(X \leq \sqrt{y}) = F_X(\sqrt{y}).$$

By differentiating and using the chain rule, we have

$$f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) = \frac{1}{2y^{1.5}}.$$

An intuitive explanation for the monotonic case



$$\begin{aligned}
 & y = g(x) \quad \delta_2 \approx \delta, \frac{\partial g}{\partial x}(x) \\
 & x = h(y) \quad \delta_1 \approx \delta_2 \cdot \frac{\partial h}{\partial y}(y) \quad \textcircled{R} \\
 & f_Y(y) \approx \mathbb{P}(y \leq Y \leq y + \delta_2) = \mathbb{P}(x \leq X \leq x + \delta_1) \\
 & \approx f_X(x) \delta_1 \approx f_X(x) \delta_2 \frac{\partial h}{\partial y}(y) \\
 & f_Y(y) = f_X(x) \frac{\partial h}{\partial y}(y) \\
 & = f_X(h(y)) \frac{\partial h}{\partial y}(y)
 \end{aligned}$$

Exercise: Using the formula for the monotonic case

5/6 points (graded)

The random variable X is exponential with parameter $\lambda = 1$. The random variable Y is defined by $Y = g(X) = 1/(1 + X)$.

a) The inverse function h , for which $h(g(x)) = x$, is of the form $ay^b + c$. Find a , b , and c .

$a =$	1	✓ Answer: 1
$b =$	-1	✓ Answer: -1
$c =$	-1	✓ Answer: -1

b) For $y \in (0, 1]$, the PDF of Y is of the form $f_Y(y) = y^a e^{(b/y)+c}$. Find a , b , and c .

$a =$	0	✗ Answer: -2
$b =$	-1	✓ Answer: -1
$c =$	1	✓ Answer: 1

Solution:

a) If x and y obey the relation $y = g(x) = 1/(1 + x)$, then $y + yx = 1$, so that

$$x = h(y) = \frac{1-y}{y} = \frac{1}{y} - 1.$$

Note that we are interested in $x \geq 0$ which restricts y to the range $(0, 1]$. Notice also that the functions g and h are monotonically decreasing on the relevant ranges of values.

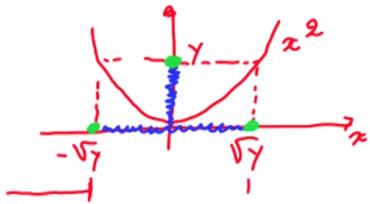
b) Note that

$$\frac{dh}{dy}(y) = -\frac{1}{y^2}.$$

Therefore,

$$f_Y(y) = f_X(h(y)) \left| \frac{dh}{dy}(y) \right| = e^{-(1/y)+1} \cdot \frac{1}{y^2}.$$

A nonmonotonic example: $Y = X^2$



- The discrete case:

$$p_Y(9) = P(x=3) + P(x=-3)$$

$$p_Y(y) = P_x(\sqrt{y}) + P_x(-\sqrt{y})$$

- The continuous case: $y \geq 0$

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) = P(|X| \leq \sqrt{y}) = P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \quad f_Y(y) = f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} + f_X(-\sqrt{y}) \frac{\cancel{-1}}{2\sqrt{y}} \end{aligned}$$

Exercise: Nonmonotonic functions

4/4 points (graded)

Suppose that X is a continuous random variable and that $Y = X^4$. Then, for $y \geq 0$, we have

$$f_Y(y) = ay^b f_X(-cy^d) + ay^b f_X(cy^d),$$

for some a, b, d , and some $c > 0$. Find a, b, c , and d .

$a =$	1/4	✓ Answer: 0.25
$b =$	-3/4	✓ Answer: -0.75
$c =$	1	✓ Answer: 1
$d =$	1/4	✓ Answer: 0.25

Solution:

We have, for $y \geq 0$,

$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(X^4 \leq y) = \mathbf{P}(-y^{1/4} \leq X \leq y^{1/4}) = F_X(y^{1/4}) - F_X(-y^{1/4}).$$

By differentiating, and using also the chain rule, we obtain

$$f_Y(y) = f_X(y^{1/4}) \cdot \frac{1}{4} \cdot y^{-3/4} + f_X(-y^{1/4}) \cdot \frac{1}{4} \cdot y^{-3/4}.$$

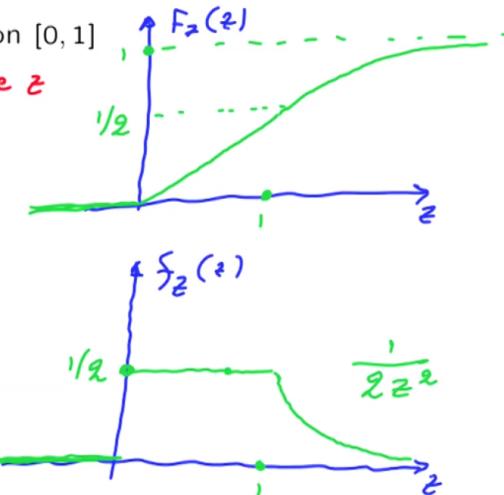
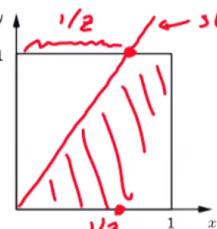
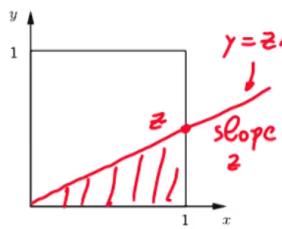
Therefore, $a = 1/4$, $b = -3/4$, $c = 1$, and $d = 1/4$.

A function of multiple r.v.'s: $Z = g(X, Y)$

$Y = g(x)$

- Same methodology: find CDF of Z

- Let $Z = Y/X$; X, Y independent, uniform on $[0, 1]$



$$\begin{aligned} F_Z(z) &= P\left(\frac{Y}{X} \leq z\right) = 0, \quad z < 0 \\ &= \frac{1}{2} \cdot z, \quad 0 \leq z \leq 1 \\ &= 1 - \frac{1}{2z}, \quad z > 1 \end{aligned}$$

Exercise: A function of multiple r.v.'s

1/2 points (graded)

Suppose that X and Y are described by a joint PDF which is uniform inside the unit circle, that is, the set of points that satisfy $x^2 + y^2 \leq 1$. In particular, the joint PDF takes the value of $1/\pi$ on the unit circle. Let $Z = \sqrt{X^2 + Y^2}$, which is the distance of the outcome (X, Y) from the origin. The PDF of Z , for $z \in [0, 1]$, takes the form $f_Z(z) = az^b$. Find a and b .

$a =$	0
$b =$	1

✗ Answer: 2

✓ Answer: 1

Solution:

Note that the set of points that satisfy $x^2 + y^2 \leq z^2$ is a circle of radius z , has area πz^2 , and probability z^2 . Therefore,

$$F_Z(z) = P(Z \leq z) = P(X^2 + Y^2 \leq z^2) = z^2,$$

from which it follows that $f_Z(z) = 2z$.

Sums of independent random variables; Covariance and Correlation

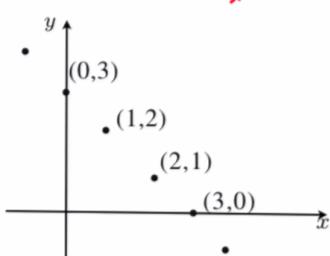
The distribution of $X + Y$: the discrete case

- $Z = X + Y$; X, Y independent, discrete

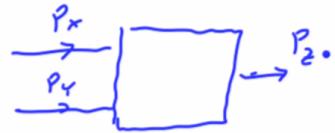
$g(x,y)$ known PMFs

$$p_Z(3) = \dots + p(x=0, y=3) + p(x=1, y=2) + \dots$$

$$= \dots + p_x(0) p_y(3) + p_x(1) p_y(2) + \dots$$



$$p_Z(z) = \sum_x p_X(x) p_Y(z-x)$$

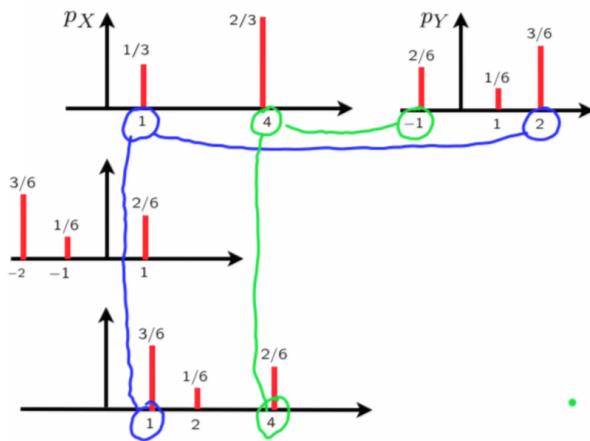


$$\begin{aligned} p_Z(z) &= \sum_x P(X=x, Y=z-x) \\ &= \sum_x p_x(x) p_y(z-x) \end{aligned}$$

Convolution formula takes in PMFs of x and y and spits out PMF of z

Discrete convolution mechanics

$$p_Z(z) = \sum_x p_X(x) p_Y(z-x)$$



- To find $p_Z(3)$:

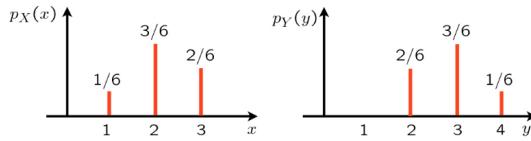
- Flip (horizontally) the PMF of Y
- Put it underneath the PMF of X
- Right-shift the flipped PMF by 3
- Cross-multiply and add
- Repeat for other values of z

Shift by value of z that we are trying to obtain

Exercise: Discrete convolution

1/1 point (graded)

The random variables X and Y are independent and have the PMFs shown in this diagram.



The probability that $X + Y = 6$ is: 1/4 ✓ Answer: 0.25

(Although you can find the answer by inspection, try to use the flip-and-shift graphical method.)

Solution:

We flip the PMF of Y to obtain a PMF on the set $\{-4, -3, -2\}$. We shift it to the right by 6 and place it underneath the PMF of X . By multiplying the probabilities that are on top of each other in the resulting diagram, we obtain

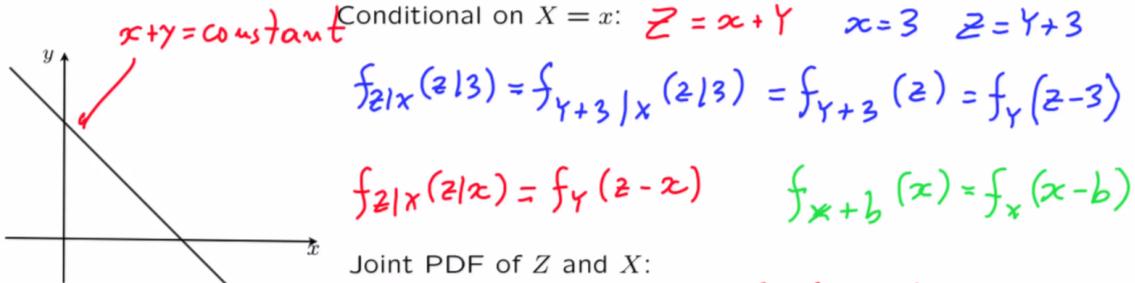
$$p_{X+Y}(6) = \frac{1}{6} \cdot \frac{3}{6} + \frac{3}{6} \cdot \frac{2}{6} = \frac{9}{36} = 1/4.$$

The distribution of $X + Y$: the continuous case

- $Z = X + Y$; X, Y independent, continuous known PDFs

$$p_Z(z) = \sum_x p_X(x) p_Y(z-x)$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$



- Same mechanics as in discrete case (flip, shift, etc.)

Exercise: Continuous convolution

2/2 points (graded)

When calculating the convolution of two PDFs, one must be careful to use the appropriate limits of integration. Suppose that X and Y are nonnegative random variables. In particular, $f_X(x)$ is equal to some positive function $h_X(x)$ for $x \geq 0$ and is zero for $x < 0$. Similarly, $f_Y(y)$ is equal to some positive function $h_Y(y)$ for $y \geq 0$, and is zero for $y < 0$. Then, the convolution integral $\int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$ is of the form

$$\int_a^b h_X(x) h_Y(z-x) dx,$$

for suitable choices of a and b determined by z . Fix some $z \geq 0$. Find a and b . (Your answer can be an algebraic function of z .)

$a =$ ✓ Answer: 0

0

$b =$ ✓ Answer: z

z

Solution:

The integrand is equal to $h_X(x) h_Y(z-x)$ only for those choices of x for which the arguments of the functions h_X and h_Y are nonnegative; that is, when $x \geq 0$ and $z-x \geq 0$, which yields $0 \leq x \leq z$. Thus, we should only integrate from 0 to z .

Graphically, the PDF of X extends from 0 to ∞ . Also, when we flip the PDF of Y , the resulting PDF extends from $-\infty$ to 0, and when we shift it to the right by z , it will extend from $-\infty$ to z . Thus the two PDFs that we need to multiply in the convolution integral overlap only for values from 0 to z .

The sum of independent normal r.v.'s

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$

- $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$, independent $Z = X + Y$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-(x-\mu_x)^2/2\sigma_x^2} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-(y-\mu_y)^2/2\sigma_y^2}$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left\{-\frac{(z-x-\mu_y)^2}{2\sigma_y^2}\right\} dx$$

$$(\text{algebra}) = \frac{1}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} \exp\left\{-\frac{(z-\mu_x-\mu_y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right\} \quad N(\mu_z, \sigma_z^2)$$

$X + Y + W$

The sum of finitely many independent normals is normal

Exercise: Sum of normals

3/3 points (graded)

Let X and Y be independent normal random variables.

a) Is $2X - 4$ always normal?

True

✓ Answer: True

b) Is $3X - 4Y$ always normal?

True

✓ Answer: True

c) Is $X^2 + Y$ always normal?

False

✓ Answer: False

Solution:

a) This is a fact that we are already familiar with: a linear function of a normal random variable is normal.

b) Since X and Y are independent and normal, the random variables $3X$ and $-4Y$ are also independent and normal. Since the sum of independent normals is normal, it follows that $3X - 4Y$ is normal.

c) There is no reason for this to be the case. To see this, consider an extreme case where $Y = 0$ (a degenerate case of a normal). Then the random variable $X^2 + Y$ is nonnegative, which is incompatible with having a normal distribution.

E[XY] based on how many data points are in each quadrant (assume same probability for all points)

Covariance tells us how random variables move together

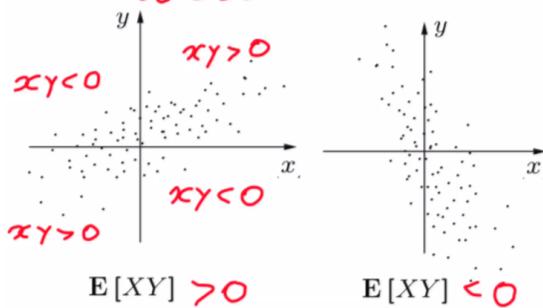
In independent case the covariance is 0

Covariance

- Zero-mean, discrete X and Y

- if independent: $E[XY] =$

$$= E[X]E[Y] = 0$$

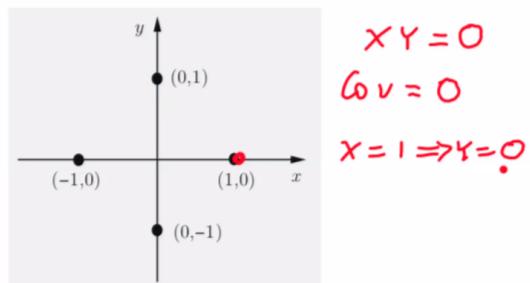


Definition for general case:

$$\text{cov}(X, Y) = E[(X - E[X]) \cdot (Y - E[Y])]$$

$$\text{and } 0 = E[(X - E[X])] E[Y - E[Y]]$$

- independent $\Rightarrow \text{cov}(X, Y) = 0$
(converse is not true)



In example given the covariance is also 0 but the r.v's are dependent as we know what y/x will be given x/y

Exercise: Covariance calculation

1/1 point (graded)

Suppose that X , Y , and Z are independent random variables with unit variance. Furthermore, $E[X] = 0$ and $E[Y] = E[Z] = 2$. Then,

$$\text{Cov}(XY, XZ) = \boxed{4} \quad \checkmark \text{ Answer: 4}$$

Solution:

Because of independence and the zero-mean assumption, it follows that $E[XY] = E[X] \cdot E[Y] = 0$ and similarly, $E[XZ] = 0$. Thus,

$$\text{Cov}(XY, XZ) = E[XYXZ] = E[X^2YZ] = E[X^2] \cdot E[Y] \cdot E[Z] = \text{Var}(X) \cdot E[Y] \cdot E[Z] = 4.$$

Covariance properties

$$\text{cov}(X, X) = E[(X - E[X])^2]$$

= $\text{var}(x) = E[X^2] - (E[X])^2$

$$\begin{aligned} \text{cov}(aX + b, Y) &= \\ (\text{assume } 0 \text{ means}) \quad &= E[(ax+b)y] = aE[xy] + bE[y] \\ &= a \cdot \text{cov}(x, y) \\ \text{cov}(X, Y + Z) &= E[X(Y+Z)] \\ &= E[XY] + E[XZ] = \text{cov}(x, y) + \text{cov}(x, z) \end{aligned}$$

$$\text{cov}(X, Y) = E[(\underbrace{X - E[X]}_{= E[X]Y} \cdot \underbrace{Y - E[Y]}_{= E[X]E[Y]})]$$

$$\begin{aligned} &= E[XY] - E[X]E[Y] \\ &\quad - E[E[X]Y] + E[E[X]E[Y]] \\ &= E[XY] - E[X]E[Y] \\ &\quad - E[X]E[Y] + E[X]E[Y] \end{aligned}$$

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

The covariances behave linearly

a) Is it true that $\text{Cov}(X, Y) = \text{Cov}(Y, X)$?

True ✓ Answer: True

b) Find the value of a in the relation $\text{Cov}(2X, -3Y + 2) = a \cdot \text{Cov}(X, Y)$.

$a =$ -6 ✓ Answer: -6

c) Suppose that X , Y , and Z are independent, with a common variance of 5. Then,

$\text{Cov}(2X + Y, 3X - 4Z) =$ 30 ✓ Answer: 30

Solution:

a) We have $(X - \mathbb{E}[X])(Y - \mathbb{E}[Y]) = (Y - \mathbb{E}[Y])(X - \mathbb{E}[X])$, and after taking expectations we obtain $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.

b) We have argued that $\text{Cov}(aX + b, Y) = a \cdot \text{Cov}(X, Y)$. Note that by symmetry, we also have $\text{Cov}(X, aY + b) = a \cdot \text{Cov}(X, Y)$. By using these relations,

$$\text{Cov}(2X, -3Y + 2) = 2 \cdot \text{Cov}(X, -3Y + 2) = 2 \cdot (-3) \cdot \text{Cov}(X, Y) = -6 \text{Cov}(X, Y).$$

c) Using linearity,

$$\begin{aligned}\text{Cov}(2X + Y, 3X - 4Z) &= \text{Cov}(2X + Y, 3X) + \text{Cov}(2X + Y, -4Z) \\ &= \text{Cov}(2X, 3X) + \text{Cov}(Y, 3X) + \text{Cov}(2X, -4Z) + \text{Cov}(Y, -4Z) \\ &= 6 \text{Var}(X) + 0 + 0 + 0 = 30,\end{aligned}$$

where the zeros are obtained because independent random variables have zero covariance.

The variance of a sum of random variables

$$\begin{aligned}\text{var}(X_1 + X_2) &= E[(x_1 + x_2 - E[x_1 + x_2])^2] \\ &= E[((x_1 - E[x_1]) + (x_2 - E[x_2]))^2] \\ &= E[(x_1 - E[x_1])^2 + (x_2 - E[x_2])^2 \\ &\quad + 2(x_1 - E[x_1])(x_2 - E[x_2])] \\ &= \text{var}(x_1) + \text{var}(x_2) + 2 \text{cov}(x_1, x_2).\end{aligned}$$

The variance of a sum of random variables

$$\text{var}(X_1 + X_2) = \text{var}(X_1) + \text{var}(X_2) + 2 \text{cov}(X_1, X_2)$$

$$\begin{aligned} \text{var}(X_1 + \dots + X_n) &= E[(X_1 + \dots + X_n)^2] \\ (\text{assume 0 means}) &= E\left[\sum_{i=1}^n X_i^2 + \sum_{\substack{i=1, \dots, n \\ j=1, \dots, n \\ i \neq j}} X_i X_j\right] \\ &= \sum_i \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \end{aligned}$$

$$\boxed{\text{var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{var}(X_i) + \sum_{\{(i,j): i \neq j\}} \text{cov}(X_i, X_j)}$$

Also true for non zero means, just to simplify this example

Exercise: The variance of a sum

1/1 point (graded)

The random variables X_1, \dots, X_8 satisfy $E[X_i] = 1$ and $\text{Var}(X_i) = 4$ for $i = 1, 2, \dots, 8$. Also, for $i \neq j$, $E[X_i X_j] = 3$. Then,

$$\text{Var}(X_1 + \dots + X_8) = \boxed{144} \quad \checkmark \text{ Answer: 144}$$

Solution:

For $i \neq j$, we have $\text{Cov}(X_i, X_j) = E[X_i X_j] - E[X_i] \cdot E[X_j] = 3 - 1 = 2$. Thus,

$$\text{Var}(X_1 + \dots + X_8) = 8 \cdot \text{Var}(X_1) + 56 \cdot \text{Cov}(X_1, X_2) = 32 + 112 = 144.$$

The Correlation coefficient

- Dimensionless version of covariance:

$$-1 \leq \rho \leq 1$$

$$\begin{aligned}\rho(X, Y) &= E\left[\frac{(X - E[X])}{\sigma_X} \cdot \frac{(Y - E[Y])}{\sigma_Y}\right] \\ &= \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}\end{aligned}$$

- Measure of the degree of "association" between X and Y
- Independent $\Rightarrow \rho = 0$, "uncorrelated" (converse is not true)
- $|\rho| = 1 \Leftrightarrow (X - E[X]) = c(Y - E[Y])$ (linearly related)
- $\text{cov}(aX + b, Y) = a \cdot \text{cov}(X, Y) \Rightarrow \rho(aX + b, Y) = \frac{a \cdot \text{cov}(X, Y)}{|a| \sigma_X \sigma_Y} = \frac{\text{sign}(a)}{\rho(X, Y)}$

Doesn't depend on the type of units we use for each of the random variables

$\text{sign}(a)$ is a function that assumes the values 1 or -1 depending on whether a is positive or negative:

$$\text{sign}(a) = \{1 \text{ if } a > 0, -1 \text{ if } a < 0\}$$

Exercise: Correlation coefficient

0/1 point (graded)

It is known that for a standard normal random variable X , we have $E[X^3] = 0$, $E[X^4] = 3$, $E[X^5] = 0$, $E[X^6] = 15$. Find the correlation coefficient between X and X^3 . Enter your answer as a number.

3

✗ Answer: 0.77460

Solution:

Since $E[X] = E[X^3] = 0$, we have $\text{Cov}(X, X^3) = E[X \cdot X^3] = E[X^4] = 3$. Furthermore, since $\text{Var}(X) = 1$ and $\text{Var}(X^3) = E[X^6] = 15$, we obtain

$$\rho(X, X^3) = \frac{3}{\sqrt{1} \cdot \sqrt{15}} = \sqrt{3/5}.$$

Interestingly, even though the random variables are strongly dependent (the value of one determines the value of the other), the value of the correlation coefficient is moderate.

Proof of key properties of the correlation coefficient

$$\rho(X, Y) = E\left[\frac{(X - E[X])}{\sigma_X} \cdot \frac{(Y - E[Y])}{\sigma_Y}\right]$$

$$-1 \leq \rho \leq 1$$

- Assume, for simplicity, zero means and unit variances, so that $\rho(X, Y) = E[XY]$

$$E[(X - \rho Y)^2] = E[X^2] - 2\rho E[XY] + \rho^2 E[Y^2]$$

$$0 \leq E[X^2] - 2\rho E[XY] + \rho^2 E[Y^2] = 1 - 2\rho^2 + \rho^2 = 1 - \rho^2 \quad 1 - \rho^2 \geq 0 \Rightarrow \rho^2 \leq 1$$

If $|\rho| = 1$, then $X = \rho Y \Rightarrow X = Y \text{ or } X = -Y$

Interpreting the correlation coefficient

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

- Association does not imply causation or influence

X : math aptitude

Y : musical ability

- Correlation often reflects underlying, common, hidden factor

- Assume, Z, V, W are independent

$$\rho(X, Y) = \frac{1}{\sqrt{2} \cdot \sqrt{2}} = \frac{1}{2}$$

$$X = \underline{Z} + V \quad Y = \underline{Z} + W$$

Assume, for simplicity, that Z, V, W have zero means, unit variances

$$\text{var}(X) = \text{var}(Z) + \text{var}(V) = 2 \Rightarrow \sigma_Z = \sqrt{2} \quad \sigma_V = \sqrt{2}$$

$$\begin{aligned} \text{cov}(X, Y) &= E[(Z+V)(Z+W)] = E[Z^2] + E[ZW] + E[VZ] + E[VW] \\ &= 1 + 0 + 0 + 0 \end{aligned}$$

Exercise: Correlation properties

5/6 points (graded)

As in the preceding example, let Z , V , and W be independent random variables with mean 0 and variance 1, and let $X = Z + V$ and $Y = Z + W$. We have found that $\rho(X, Y) = 1/2$.

a) It follows that:

$$\rho(X, -Y) = \boxed{-1/2}$$

✓ Answer: -0.5

$$\rho(-X, -Y) = \boxed{1/2}$$

✓ Answer: 0.5

b) Suppose that X and Y are measured in dollars. Let X' and Y' be the same random variables, but measured in cents, so that $X' = 100X$ and $Y' = 100Y$. Then,

$$\rho(X', Y') = \boxed{1/2}$$

✓ Answer: 0.5

c) Suppose now that $\tilde{X} = 3Z + 3V + 3$ and $\tilde{Y} = -2Z - 2W$. Then

$$\rho(\tilde{X}, \tilde{Y}) = \boxed{0}$$

✗ Answer: -0.5

d) Suppose now that the variance of Z is replaced by a very large number. Then

$$\rho(X, Y) \text{ is close to } \boxed{1}$$

✓ Answer: 1

e) Alternatively, suppose that the variance of Z is close to zero. Then

$$\rho(X, Y) \text{ is close to } \boxed{0}$$

✓ Answer: 0

Solution:

We saw that a linear transformation $x \mapsto ax + b$ of a random variable does not change the value of the correlation coefficient, except for a possible sign change if the coefficient a is negative. Note that in the case of $\rho(-X, -Y)$, we have two sign changes, hence no sign change.

For the last two parts, if Z has a very large variance, then the terms V and W become insignificant, and $\rho(X, Y) \approx \rho(Z, Z) = 1$. And if Z has very small variance, then X and Y are approximately independent, so that $\rho(-X, -Y) \approx 0$. (These conclusions can also be justified by an exact calculation.)

Correlations matter...

- A real-estate investment company invests \$10M in each of 10 states. At each state i , the return on its investment is a random variable X_i , with mean 1 and standard deviation 1.3 (in millions).

$$\text{var}(X_1 + \dots + X_{10}) = \sum_{i=1}^{10} \text{var}(X_i) + \sum_{\{(i,j): i \neq j\}} \text{cov}(X_i, X_j)$$

$E[X_1 + \dots + X_{10}] = 10$

- If the X_i are uncorrelated, then:

$$\text{var}(X_1 + \dots + X_{10}) = 10 \cdot (1.3)^2 = 16.9 \quad \sigma(X_1 + \dots + X_{10}) = 4.1$$

- If for $i \neq j$, $\rho(X_i, X_j) = 0.9$:

$$\text{cov}(X_i, X_j) = \rho \sigma_{X_i} \sigma_{X_j} = 0.9 \times 1.3 \times 1.3 = 1.52$$

$$\text{var}(X_1 + \dots + X_{10}) = 10 \cdot (1.3)^2 + 90 \cdot 1.52 = 154$$

$$\sigma(X_1 + \dots + X_{10}) = 12.4$$

In uncorrelated case then positive profit is likely

But in correlated case there's a high probability that there will be a loss

This is kind of what happened during the housing crisis since many investors thought they were making diverse investments but they were all reliant on similar factors