6 Jointly distributed r.v.s

6.1 Joint dist. functions

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

joint distribution function

marginal dist. funcs :

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y)$$

$$F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x,y)$$

If both X and Y are discrete:

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

joint mass function

marginal mass funcs :

$$p_X(x) = \sum_{y \in R(Y)} p_{X,Y}(x,y)$$

$$p_Y(y) = \sum_{x \in R(X)} p_{X,Y}(x,y)$$

X and Y are *jointly continuous* if they have a *joint density function* $f_{X,Y}(x,y)$ satisfying

$$P((X,Y) \in D) = \iint_D f_{X,Y}(x,y)\, dA,$$

whenever $D \subseteq \mathbb{R}^2$.

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)\, dt\, ds$$

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

$$P(x \le X \le x+\Delta x,\ y \le Y \le y+\Delta y) \approx f_{X,Y}(x,y)\,\Delta x \Delta y$$

marginal densities:

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

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

Example ~~1c~~ 1d

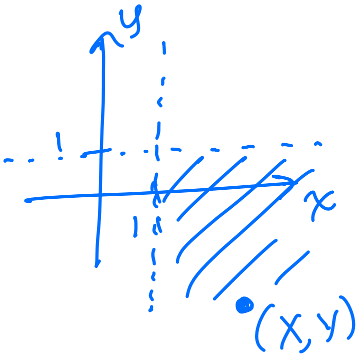
The joint density function of X and Y is given by

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

Compute (a) $P\{X > 1, Y < 1\}$, (b) $P\{X < Y\}$, and (c) $P\{X < a\}$.

(a) $$P(X>1, Y<1) = \int_1^\infty \int_{-\infty}^1 f(x,y)\,dy\,dx$$

$$\boxed{\int_{-\infty}^\infty \int_{-\infty}^\infty f(x,y)\,dx\,dy = 1}$$



$$= \int_1^\infty \int_0^1 2e^{-x} e^{-2y}\,dy\,dx$$

$$= \int_1^\infty 2e^{-x}\left(-\frac{1}{2}e^{-2y}\Big|_{y=0}^{y=1}\right)dx$$

$$= \int_1^\infty 2e^{-x}\left(-\frac{1}{2}e^{-2} + \frac{1}{2}\right)dx$$

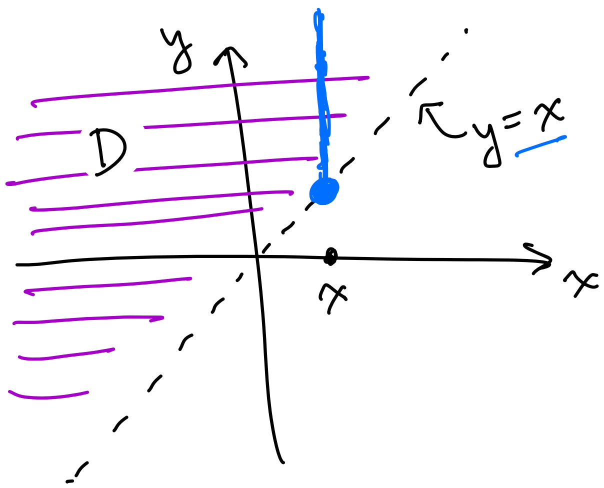
$$= (1-e^{-2})\int_1^\infty e^{-x}\,dx$$

$$= (1-e^{-2})\left(-e^{-x}\Big|_{x=1}^{x=\infty}\right)$$

$$= (1-e^{-2})\left(0-(-e^{-1})\right)$$

$$= \boxed{e^{-1}(1-e^{-2})}$$

(b) $P(X<Y) = ?$



$$D = \{(x,y) : x<y\}$$

$$\{X<Y\} = \{(X,Y) \in D\}$$

$$P(X<Y) = P((X,Y)\in D) = \iint_D f(x,y)\,dA$$

$$= \int_{-\infty}^{\infty} \int_{x}^{\infty} f(x,y)\,dy\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{y} f(x,y)\,dx\,dy$$

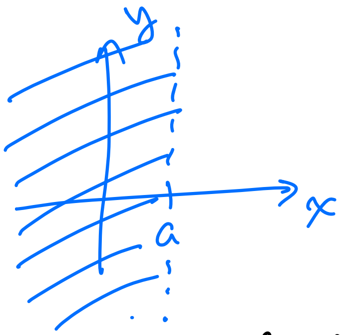
$$= \int_0^{\infty}\int_x^{\infty} 2e^{-x}e^{-2y}\,dy\,dx$$

$$= \int_0^{\infty} 2e^{-x}\left(-\frac{1}{2}e^{-2y}\Big|_{y=x}^{y=\infty}\right)dx$$

$$= \int_0^{\infty} 2e^{-x}\left(0 - \left(-\frac{1}{2}e^{-2x}\right)\right)dx$$

$$= \int_0^{\infty} e^{-3x}\,dx = \boxed{\frac{1}{3}}$$

(c) $P(X<a) = P(X<a, Y<\infty)$



$$= \int_{-\infty}^{a}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx$$

$$= \int_0^a \int_0^{\infty} 2e^{-x}e^{-2y}\,dy\,dx$$

$$= \int_0^a 2e^{-x}\left(-\frac{1}{2}e^{-2y}\Big|_{y=0}^{y=\infty}\right)dx$$

$$= \int_0^a 2e^{-x}\left(0-\left(-\frac{1}{2}\right)\right)dx$$

$$= \int_0^a e^{-x}\,dx = \left(-e^{-x}\right)\Big|_{x=0}^{x=a}$$

$$= -e^{-a} - (-1) = \boxed{1-e^{-a}}$$

Example ~~1d~~ 1e

Consider a circle of radius R, and suppose that a point within the circle is randomly chosen in such a manner that all regions within the circle of equal area are equally likely to contain the point. (In other words, the point is uniformly distributed within the circle.) If we let the center of the circle denote the origin and define X and Y to be the coordinates of the point chosen (Figure 1), then, since (X, Y) is equally likely to be near each point in the circle, it follows that the joint density function of X and Y is given by

$$f(x, y) = \begin{cases} c & \text{if } x^2 + y^2 \le R^2 \\ 0 & \text{if } x^2 + y^2 > R^2 \end{cases}$$

for some value of c.

(a) Determine c.

(b) Find the marginal density functions of X and Y.

(c) Compute the probability that D, the distance from the origin of the point selected, is less than or equal to a.

(d) Find $E[D]$.

(a) $1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy$

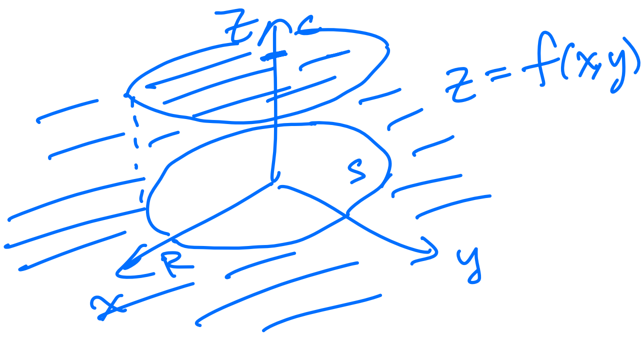
$= \iint_S c\,dA$ ← S: disk of radius R, centered at origin

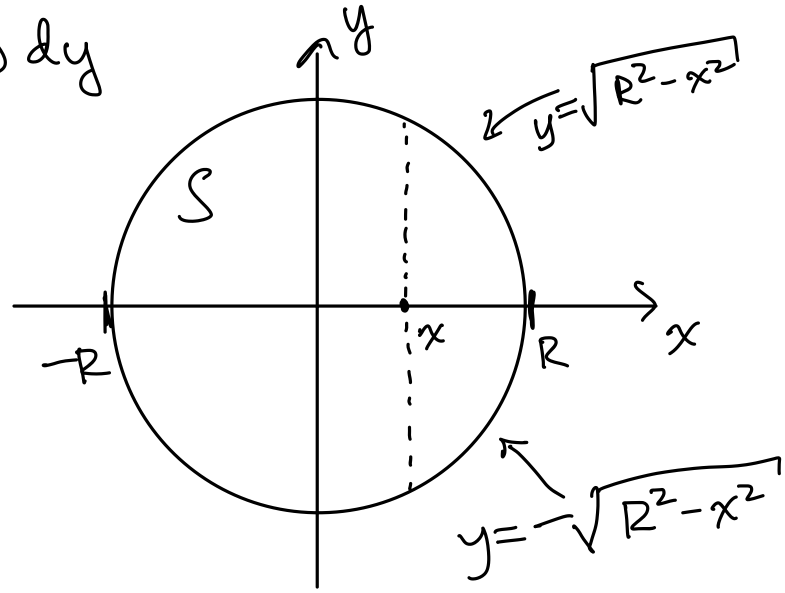
$= c \cdot \text{area}(S)$

$= c \cdot \pi R^2$

$\Rightarrow \quad c = \boxed{\dfrac{1}{\pi R^2}}$

(b) $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$





for $-R \le x \le R$:

$$f_X(x) = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} c\,dy = 2c\sqrt{R^2-x^2}$$

$$= \frac{2}{\pi R^2}\sqrt{R^2-x^2}$$

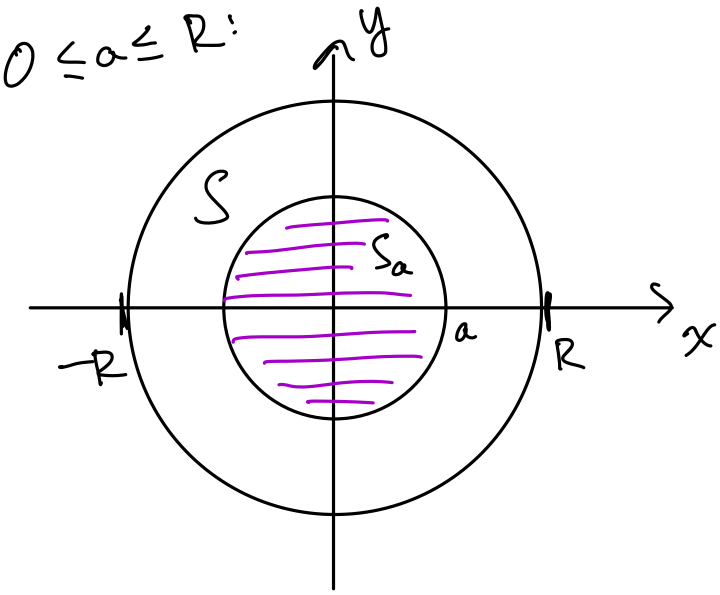
$$\boxed{f_X(x) = \begin{cases} \frac{2}{\pi R^2}\sqrt{R^2-x^2} & \text{if } -R \le x \le R, \\ 0 & \text{otherwise} \end{cases}}$$

By symmetry:

$$\boxed{f_Y(y) = \begin{cases} \frac{2}{\pi R^2}\sqrt{R^2-y^2} & \text{if } -R \le y \le R, \\ 0 & \text{otherwise} \end{cases}}$$

(c) $D = \sqrt{X^2 + Y^2}$

for $0 \le a \le R$:



$\{D \le a\} = \{(X, Y) \in S_a\}$

$$P(D \le a) = \iint_{S_a} f(x,y)\, dA$$

$$= \iint_{S_a} c\, dA$$

$$= c \cdot \text{area}(S_a)$$

$$= \frac{1}{\pi R^2} \cdot \pi a^2 = \frac{a^2}{R^2}$$

$$\boxed{P(D \le a) = \begin{cases} 0 & \text{if } a < 0, \\ \frac{a^2}{R^2} & \text{if } 0 \le a \le R, \\ 1 & \text{if } a > R \end{cases}}$$

(d) $f_D(a) = F_D'(a)$ $\qquad (F_D(a) = P(D \leq a))$

$$= \begin{cases} 0 & \text{if } a < 0, \\ \frac{2a}{R^2} & \text{if } 0 \leq a \leq R, \\ 0 & \text{if } a > R. \end{cases}$$

$$E[D] = \int_{-\infty}^{\infty} a f_D(a)\, da$$

$$= \int_0^R a \cdot \frac{2a}{R^2}\, da$$

$$= \frac{2}{R^2} \int_0^R a^2\, da = \frac{2}{R^2} \cdot \frac{1}{3} a^3 \Big|_{a=0}^{a=R}$$

$$= \frac{2}{\cancel{R^2}} \cdot \frac{R^{\cancel{3}}}{3} = \boxed{\frac{2R}{3}}$$

6.2 Indep. r.v.s

X and Y are <u>independent</u> if

$$P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B)$$

whenever $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}$.

Extends to more than 2 r.v.s just like events.

All of the following are equivalent to independence:

- $F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$
- $p_{X,Y}(x,y) = p_X(x)\,p_Y(y)$ if X, Y discrete
- $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$ if X, Y jointly cont.

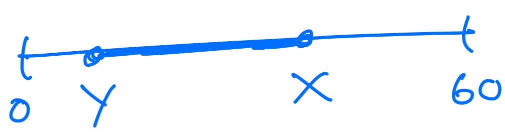
Example 2c A man and a woman decide to meet at a certain location. If each of them independently arrives at a time uniformly distributed between 12 noon and 1 P.M., find the probability that the first to arrive has to wait longer than 10 minutes.

X = # of minutes after 12:00 that the man arrives.
Y = " woman "

$X \sim \text{Unif}(0,60)$
$Y \sim \text{Unif}(0,60)$
X, Y indep.

A = "The first to arrive has to wait longer than 10 minutes."

$A = \{|X-Y| > 10\}$



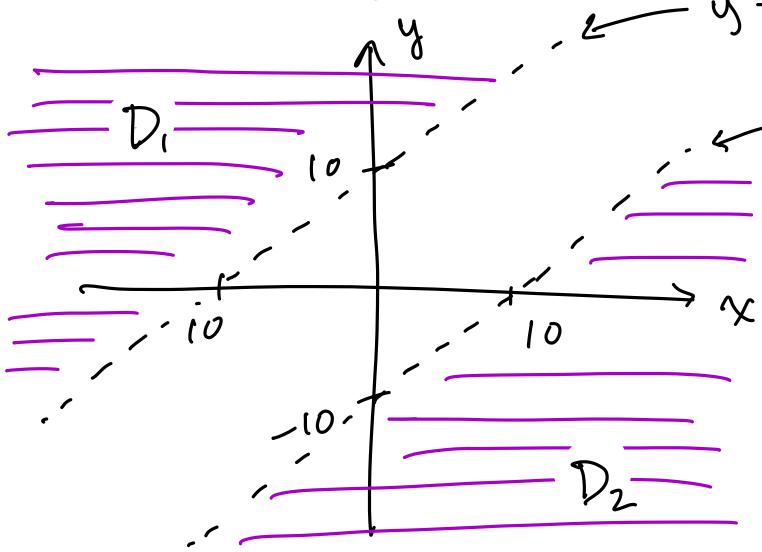
$P(A) = P(|X-Y| > 10) = ?$

$x - y > 10$ or $x - y < -10$

$D = \{(x,y) : |x-y| > 10\}$

$y = x + 10$ $(x - y = -10)$

$y = x - 10$ $(x - y = 10)$



$D = D_1 \cup D_2$

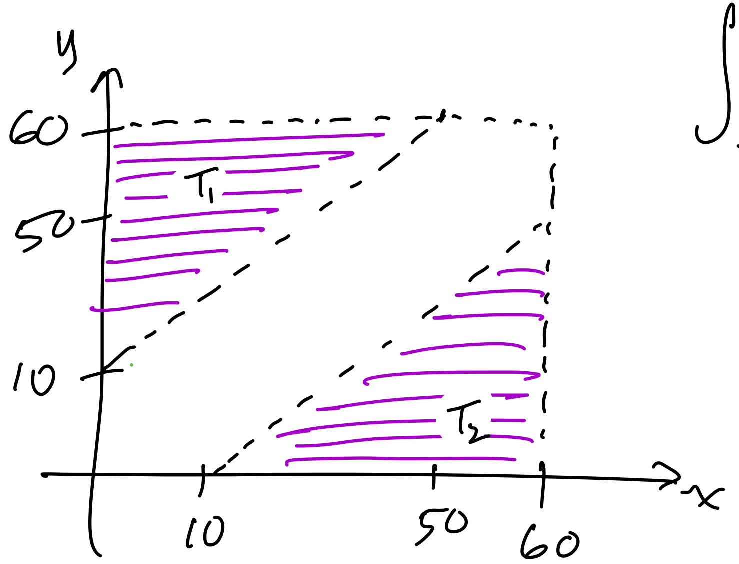
$$P(|X-Y|>10) = \iint_D \underbrace{f_{X,Y}(x,y)}_{=?}\, dA$$

$f_{X,Y}(x,y) = f_X(x) f_Y(y)$ b/c X, Y are indep.

$$f_X(x) = \begin{cases} \frac{1}{60} & \text{if } 0 < x < 60, \\ 0 & \text{otherwise.} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{60} & \text{if } 0 < y < 60, \\ 0 & \text{otherwise} \end{cases}$$

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{3600} & \text{if } 0 < x < 60,\ 0 < y < 60, \\ 0 & \text{otherwise.} \end{cases}$$



$$\iint_D f_{X,Y}(x,y)\,dA$$

$$= \iint_{T_1} \frac{1}{3600}\,dA + \iint_{T_2} \frac{1}{3600}\,dA$$

$$= \frac{1}{3600}\,\underbrace{\cancel{\text{area}(T_1)}}_{\frac{50^2}{2} = 1250} + \frac{1}{3600}\,\underbrace{\cancel{\text{area}(T_2)}}_{1250} = \frac{2500}{3600} = \boxed{\frac{25}{36}}$$

Example 2f (a) If the joint density function of X and Y is

$$f(x,y) = 6e^{-2x}e^{-3y} \qquad 0 < x < \infty,\ 0 < y < \infty$$

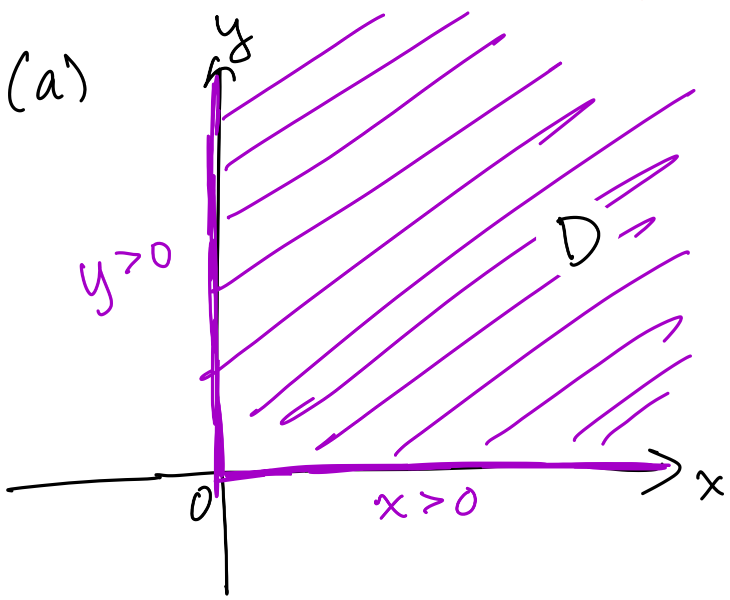
and is equal to 0 outside this region, are the random variables independent? (b) What if the joint density function is

$$f(x,y) = 24xy \qquad 0 < x < 1, 0 < y < 1,\ 0 < x + y < 1$$

and is equal to 0 otherwise?

(a') What are the marginal densities of X and Y?

(a)



$$f(x,y) = \begin{cases} 6e^{-2x}e^{-3y} & \text{if } (x,y) \in D, \\ 0 & \text{otherwise.} \end{cases}$$

This (infinite) rectangle is the product of these two intervals

If $g(x) = \begin{cases} 6e^{-2x} & \text{if } x > 0, \\ 0 & \text{if } x \le 0, \end{cases}$ and

$$h(y) = \begin{cases} e^{-3y} & \text{if } y > 0, \\ 0 & \text{if } y \le 0, \end{cases}$$

then $f(x,y) = g(x)h(y)$.

So f factors as a product of a function of x and a function of y. ← this is enough to ensure independence

Thus, X and Y are independent.

(a') g and h are not the marginal densities. To adjust them, find the correct constants:

$$\int_{-\infty}^{\infty} g(x)\,dx = \int_0^{\infty} 6e^{-2x}\,dx = -3e^{-2x}\Big|_{x=0}^{x=\infty}$$

$$= (0-(-3)) = 3$$

So $\displaystyle\int_{-\infty}^{\infty} \frac{1}{3} g(x)\,dx = 1.$

This means

$$\boxed{f_X(x) = \frac{1}{3} g(x) = \begin{cases} 2e^{-2x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}}$$

$$\int_{-\infty}^{\infty} h(y)\,dy = \int_0^{\infty} e^{-3y}\,dy = -\frac{1}{3} e^{-3y} \Big|_{y=0}^{y=\infty}$$

$$= \left(0 - \left(-\frac{1}{3}\right)\right) = \frac{1}{3}$$

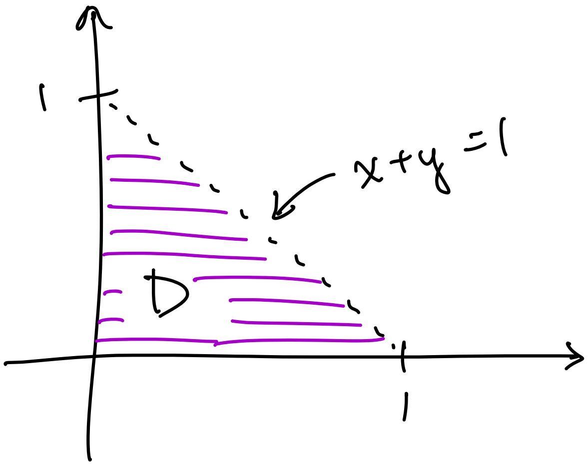
So $\displaystyle\int_{-\infty}^{\infty} 3h(y)\,dy = 1$

The two numbers in purple on this page will always be reciprocals of each other.

This means

$$\boxed{f_Y(y) = 3h(y) = \begin{cases} 3e^{-3y} & \text{if } y > 0, \\ 0 & \text{if } y \leq 0. \end{cases}}$$

(b)



$$f(x,y) = \begin{cases} 24xy & \text{if } (x,y) \in D, \\ 0 & \text{otherwise.} \end{cases}$$

The region on which f is nonzero is nonrectangular.

That means f cannot factor into a function of x and a function of y.

So X and Y are $\boxed{\text{dependent}}$.

(Any time you multiply a piecewise function of x and a piecewise function of y, their two pieces will combine to form a rectangle. If the combined region is not a rectangle, it could not have been a product of that form.)

Functions of indep. r.v.s

If $X_1, X_2, \ldots, X_7$ are indep., then

$$\underbrace{f(X_1, X_6, X_7)}_{\text{built from different } X_j\text{'s}} \text{ and } \underbrace{g(X_2, X_5)}_{} \text{ are indep.}$$

built from different X_j's

Same principle applies to any indep. collection

6.3 Sums of indep. r.v.s

If X and Y are cont. and independent, then

$$f_{X+Y}(z) = \underbrace{\int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy}_{\text{convolution of } f_X \text{ and } f_Y}$$

Example 3a **Sum of two independent uniform random variables**

If X and Y are independent random variables, both uniformly distributed on $(0, 1)$, calculate the probability density of $X + Y$.

$Z = X + Y$, $R(Z) = (0,2)$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy$$

$$f_Y(y) = \begin{cases} 1 & \text{if } 0 < y < 1, \\ 0 & \text{otherwise}. \end{cases}$$

Using indicators:

$$f_Y(y) = 1_{(0,1)}(y)$$

remember:

$$\left(1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \right)$$

$$f_X(x) = 1_{(0,1)}(x)$$

$$f_X(z-y) = 1_{(0,1)}(z-y) = 1_{(z-1,z)}(y)$$

$$\left(\begin{array}{c} 0 < z-y < 1 \\ -z < -y < 1-z \\ z > y > z-1 \end{array} \right)$$

for $0 < z < 2$:

$$f_Z(z) = \int_{-\infty}^{\infty} 1_{(z-1,z)}(y)\, 1_{(0,1)}(y)\, dy$$

$$= \int_{-\infty}^{\infty} 1_{(z-1,z)\cap(0,1)}(y)\, dy$$

$$(z-1,z) \cap (0,1) = ?$$

if $z < 1$: $(z-1,z) \cap (0,1) = (0,z)$

if $z \geq 1$: $(z-1,z) \cap (0,1) = (z-1,1)$

So...

for $0 < z < 1$:

$$f_Z(z) = \int_{-\infty}^{\infty} 1_{(0,z)}(y)\,dy = \int_0^z 1\,dy = z$$

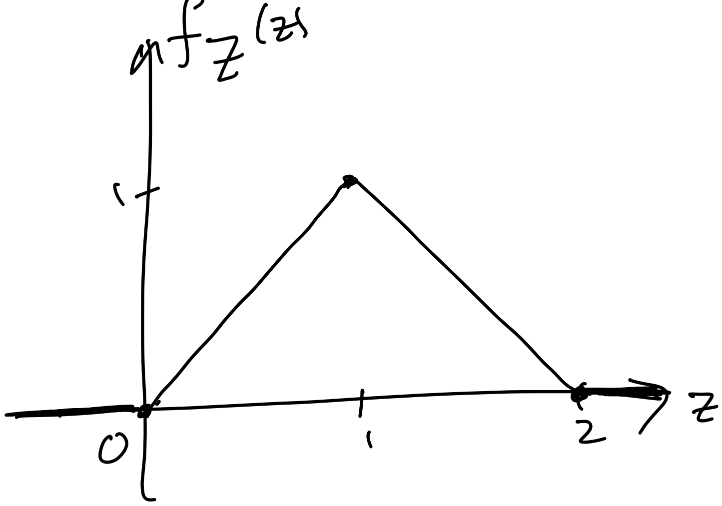
for $1 \le z < 2$:

$$f_Z(z) = \int_{-\infty}^{\infty} 1_{(z-1,1)}(y)\,dy = \int_{z-1}^{1} 1\,dy$$

$$= 1 - (z-1) = 2 - z$$

Putting it all together:

$$\boxed{f_Z(z) = \begin{cases} z & \text{if } 0 < z < 1, \\ 2 - z & \text{if } 1 \le z < 2, \\ 0 & \text{otherwise.} \end{cases}}$$



same λ

If $X \sim \text{Gamma}(\alpha, \lambda)$ and $Y \sim \text{Gamma}(\beta, \lambda)$ are indep., then $X+Y \sim \text{Gamma}(\alpha+\beta, \lambda)$.

Gamma$(1,\lambda)$

If $X_1, \dots, X_n$ are indep. $\text{Exp}(\lambda)$, then $X_1 + \dots + X_n \sim \text{Gamma}(n, \lambda)$

If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are indep., then $X+Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

Example 3d Starting at some fixed time, let $S(n)$ denote the price of a certain security at the end of n additional weeks, $n \geq 1$. A popular model for the evolution of these prices assumes that the price ratios $S(n)/S(n-1), n \geq 1$, are independent and identically distributed lognormal random variables. Assuming this model, with parameters $\mu = .0165, \sigma = .0730$, what is the probability that

(a) the price of the security increases over each of the next two weeks?

(b) the price at the end of two weeks is higher than it is today?

n: time (weeks)

$S(n)$: price of security at time n

$$Y_n = \frac{S(n)}{S(n-1)}$$

$Y_1, Y_2, Y_3, \dots$ independent

Y_n is lognormal with $\mu = 0.0165$, $\sigma = 0.0730$

That means $Y_n = e^{X_n}$, where $X_n \sim N(\mu, \sigma^2)$

(a) $P(S(1) > S(0),\ S(2) > S(1)) = ?$

(b) $P(S(2) > S(0)) = ?$

(a) $S(1) > S(0) \iff \dfrac{S(1)}{S(0)} > 1$

$$\iff Y_1 > 1$$

$$\iff e^{X_1} > 1$$

$$\iff X_1 > 0$$

For the same reasons,

$$S(2) > S(1) \iff \cdots \iff X_2 > 0$$

$P(S(1) > S(0),\ S(2) > S(1))$

$$= P(X_1 > 0,\ X_2 > 0)$$

$$= P(X_1 > 0)\, P(X_2 > 0)$$

How do we know the X_n's are independent?

$$P(X_1 > 0) = P\Big(\overbrace{\frac{X_1 - \mu}{\sigma}}^{N(0,1)} > -\frac{\mu}{\sigma}\Big) = \Phi\left(\frac{\mu}{\sigma}\right)$$

$$P(X_2 > 0) = P(X_1 > 0) = \Phi\left(\frac{\mu}{\sigma}\right)$$

$$P(S(1) > S(0),\ S(2) > S(1)) = \left(\Phi\left(\frac{\mu}{\sigma}\right)\right)^2$$

$$= \left(\Phi\left(\frac{0.0165}{0.0730}\right)\right)^2$$

$$\approx \left(\Phi(0.23)\right)^2$$

$$\approx (0.5910)^2$$

$$\approx \boxed{0.3493}$$

(b) $S(2) > S(0) \iff \dfrac{S(2)}{S(0)} > 1$

$$\iff \frac{S(2)}{S(1)} \cdot \frac{S(1)}{S(0)} > 1$$

$$\iff Y_2 Y_1 > 1$$

$$\iff e^{X_2} e^{X_1} > 1$$

$$\iff e^{X_1 + X_2} > 1$$

$$\iff X_1 + X_2 > 0$$

X_1, X_2 indep. $N(\mu, \sigma^2)$

So $X_1 + X_2 \sim N(2\mu, 2\sigma^2)$

$$X_1 + X_2 \overset{d}{=} 2\mu + \sqrt{2}\,\sigma Z \leftarrow N(0,1)$$

$$P(S(2) > S(1)) = P(X_1 + X_2 > 0)$$

$$= P(2\mu + \sqrt{2}\,\sigma Z > 0) = P\left(Z > -\frac{\overset{\sqrt{2}}{\cancel{2}}\mu}{\cancel{\sqrt{2}}\sigma}\right)$$

$$= \Phi\left(\frac{\sqrt{2}\mu}{\sigma}\right) = \Phi\left(\frac{0.0165\sqrt{2}}{0.0730}\right)$$

$$\approx \Phi(0.32) \approx \boxed{0.6255}$$

Similar results for discrete r.v.s:

same p

If $X \sim \text{Binom}(n, p)$ and $Y \sim \text{Binom}(m, p)$ are indep.

then $X + Y \sim \text{Binom}(n+m, p)$

Example 3e **Sums of independent Poisson random variables**

If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, compute the distribution of $X + Y$.

$Z = X + Y$, $R(Z) = \{0, 1, 2, \ldots\}$

For $k \in R(Z)$:

$$P(Z = k) = \sum_{j=0}^{k} P(X = k-j, Y = j)$$

$$= \sum_{j=0}^{k} P(X = k-j)\,P(Y = j)$$

$$= \sum_{j=0}^{k} e^{-\lambda_1} \cdot \frac{\lambda_1^{k-j}}{(k-j)!} \cdot e^{-\lambda_2} \cdot \frac{\lambda_2^{j}}{j!}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{j=0}^{k} \frac{1}{j!\,(k-j)!} \lambda_1^{k-j} \lambda_2^{j}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{j=0}^{k} \frac{1}{k!} \cdot \frac{k!}{j!\,(k-j)!} \lambda_1^{k-j} \lambda_2^{j}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \underbrace{\sum_{j=0}^{k} \binom{k}{j} \lambda_1^{k-j} \lambda_2^{j}}_{\text{binomial theorem}}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} (\lambda_1+\lambda_2)^k$$

So...

$$P(Z=k) = e^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^k}{k!} \quad \text{for } k=0,1,2,\ldots$$

That means

$$Z \sim \boxed{\text{Poisson}(\lambda_1+\lambda_2)}$$

HW: Ch.6: 2, 7, 8, 9, 10, 13, 15, 19,
20, 21, 22, 23, 26, 30, 37
33