We saw last time how one could graph the pdf of a Normal (μ, σ^2) distribution. However, computing areas underneath the pdf over finite (or even semi-infinite) intervals cannot be done exactly as there is no antiderivative (in closed-form) of the Normal pdf. Therefore, probabilities must be computed either numerically (eg. Simpson's rule, trapezoid rule, etc) or we must resort to tables.

At first glance you may think that tables would not be a convenient way of computing areas since we might need one such table for each possible choice of μ and σ^2. But the following fact is <u>very</u> <u>useful</u> when dealing with Normal(μ, σ^2) random variables:

<u>FACT</u>

If $X \sim$ Normal(μ, σ^2), then

$$\frac{X-\mu}{\sigma} \sim \text{Normal}(0,1)$$

the transformation $X \mapsto \frac{X-\mu}{\sigma}$ is called standardization or a z-score transformation

That is, every normal r.v. can be transformed into a standard normal r.v. by subtracting the mean and the dividing by the standard deviation.

To see why the fact is true ....

Suppose we want to compute $P(X \le x)$ for a normal r.v. X having μ and σ as its parameters. (μ and σ is known).

Then

$$P(X \le x) = \int_{-\infty}^{x} \frac{e^{-\frac{1}{2}\left\{\frac{t-\mu}{\sigma}\right\}^2}}{\sigma\sqrt{2\pi}}\, dt$$

$$= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{e^{-\frac{1}{2}\{u\}^2}}{\sqrt{2\pi}}\, du$$

make a change of varible

$u = \frac{t-\mu}{\sigma}$

$du = \frac{dt}{\sigma}$

$$= P\left(Z \le \frac{x-\mu}{\sigma}\right), \text{ where}$$

Z has pdf $\frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}$ for $-\infty < u < \infty$.

i.e., a standard normal distribution.

So $P(X \le x) = P(X-\mu \le x-\mu) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right)$

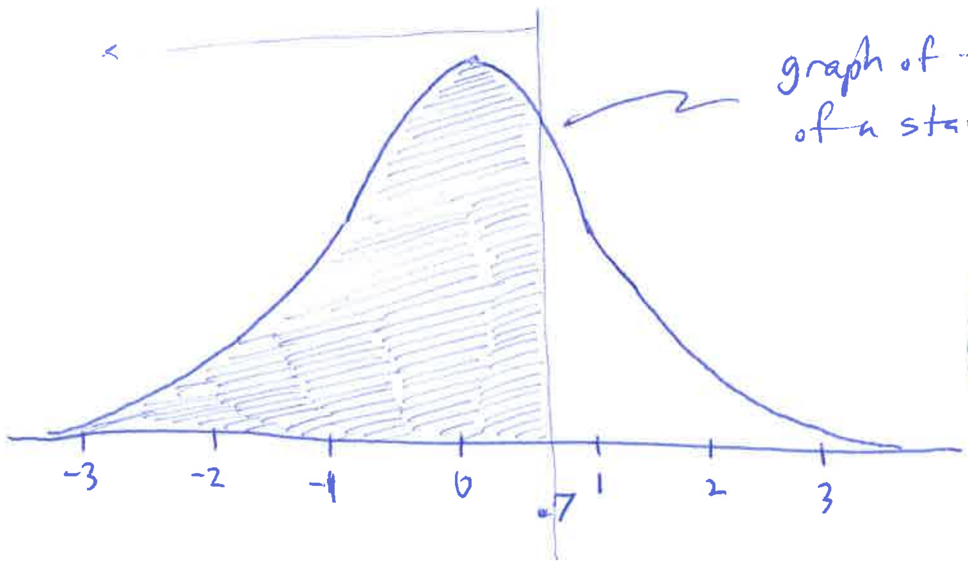
$$= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du \Rightarrow \frac{X-\mu}{\sigma} \sim \text{Normal}(0,1).$$

Using this fact, since every Normal r.v. can be transformed into a Normal r.v. having $\mu = 0$ and $\sigma = 1$, we ONLY Need ONE table to compute probabilities! Such a table is called a Standard Normal table.

The cdf of a Standard Normal r.v. gets the symbol Φ:

Special notation for cdf of a Standard Normal r.v. $Z \sim N(0,1)$

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du.$$



graph of the PDF of a standard Normal r.v.

using table on page 155 of our textbook

The shaded area above $= P(Z \leq .7) = \Phi(.7) = .7580.$

And

$$P(1.1 \leq Z \leq 2.23) = \Phi(2.23) - \Phi(1.1)$$

again using table on pg 155

$$= .9871 - .8643 = .1228.$$

Remarks about using the table on page 155

You will notice the table only gives the cdf $\Phi(z)$ for values of z ranging from

$$z = 0.00 \text{ to } z = 3.49 \text{ in increments of } .01.$$

Therefore, if we need $\Phi(1.827)$ I suggest just rounding 1.827 to 1.83 and return $\Phi(1.83)$ instead. Another possibility is to linearly interpolate the values:

$$\Phi(1.827) \approx A \cdot \Phi(1.82) + B\,\Phi(1.83)$$

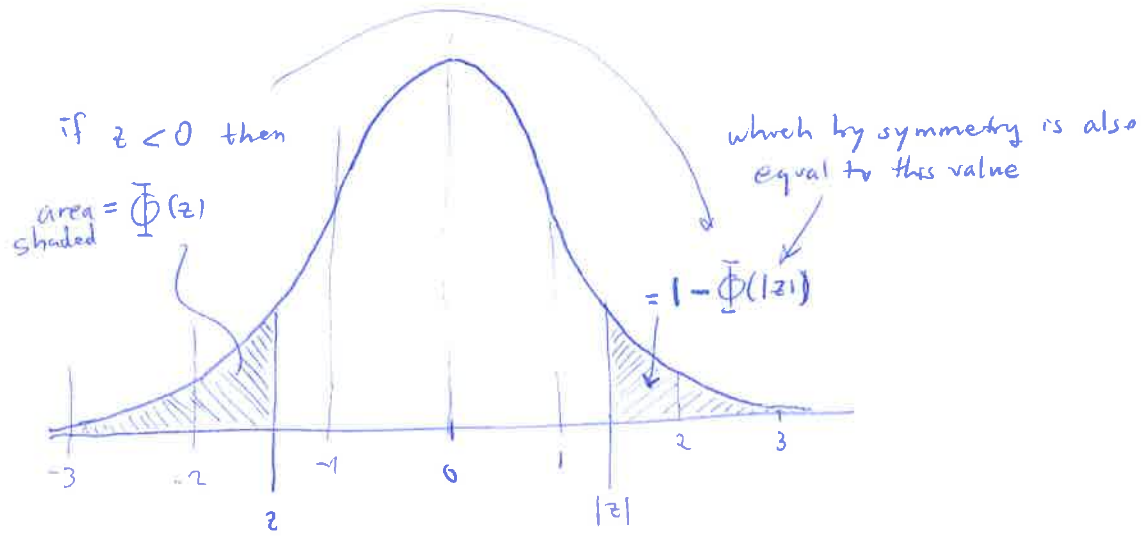
where

$$A = \frac{1.83 - 1.827}{1.83 - 1.82} = .3 \quad \text{and} \quad B = \frac{1.827 - 1.82}{1.83 - 1.82} = .7$$

In actuality $\Phi(1.827) = .982003\ldots$ while

$$.3\,\Phi(1.82) + .7\,\Phi(1.83) = .982002\ldots \text{ (fairly close!)}$$

Also, since this table only tabulates $\Phi(z)$ for $z \geq 0$ we need to use Symmetry of the PDF curve to obtain $\Phi(z)$ for $z < 0$.

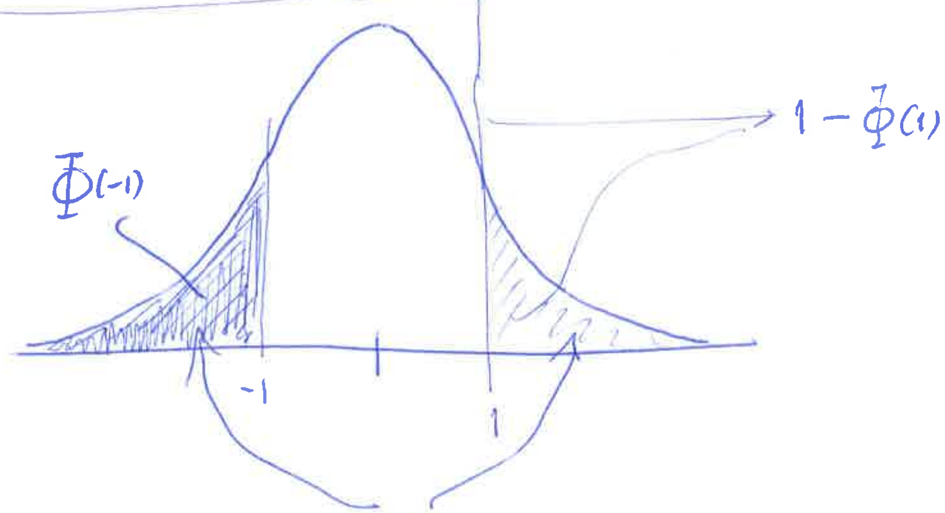


So, for example if $z>0$ then $\Phi(-z) = 1-\Phi(z)$:

<u>Example</u>

$$\Phi(-1) = 1-\Phi(1) = 1 - .8413 = .1587$$

see shading below: $\Phi(1)$ = all area under curve to left of 1.



by symmetry these two shaded areas are the <u>same</u>.

Joint pdfs, marginal pdfs, and Conditional pdfs

Since there is such a strong analogy between these concepts and those of the joint pmfs, marginal pmfs, and conditional pmfs

I will go through these ideas rather quickly.

Two r.v.s X, Y are said to be (jointly) continuous with joint pdf $f_{X,Y}(x,y)$ if for every $a \le b, c \le d$

$$P(a \le X \le b,\ c \le Y \le d) = \int_c^d \int_a^b f_{X,Y}(x,y)\, dx\, dy$$

The joint pdf $f_{X,Y}$ has the two properties

(1) $f_{X,Y}(x,y) \geq 0$ for all real x, y

(2) $\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dx\, dy = 1.$

Example

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{2x+y}{4} & \text{for } 0 \le x \le 1,\ 0 \le y \le 2 \\ 0 & \text{for other } x, y \end{cases}$$

is a joint pdf.

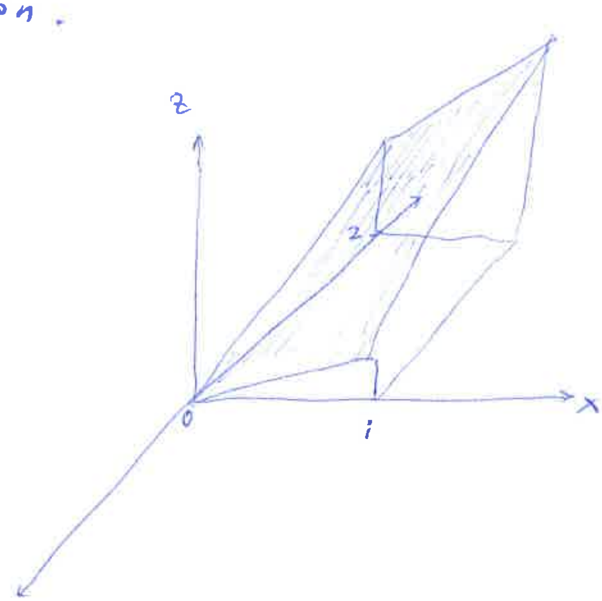
It is certainly ≥ 0 for all x, y real.

Moreover

$$\int_0^2 \left\{ \int_0^1 \frac{2x+y}{4}\, dx \right\} dy = \int_0^2 \left\{ \frac{x^2}{4} + \frac{xy}{4} \Big]_{x=0}^{x=1} \right\} dy$$

$$= \int_0^2 \frac{1+y}{4}\, dy = \frac{y}{4} + \frac{y^2}{8} \Big]_{y=0}^{y=2} = \frac{1}{2} + \frac{1}{2} = 1.$$

and we have normalization.



Recall

The marginal pdf of X (and of Y) are given as

integrate out y for fixed x.

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

integrate out x for fixed y.

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

the Conditional pdf of X given $Y=y$ is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

cannot be 0 density so this function is defined only on the essential domain of f_Y.

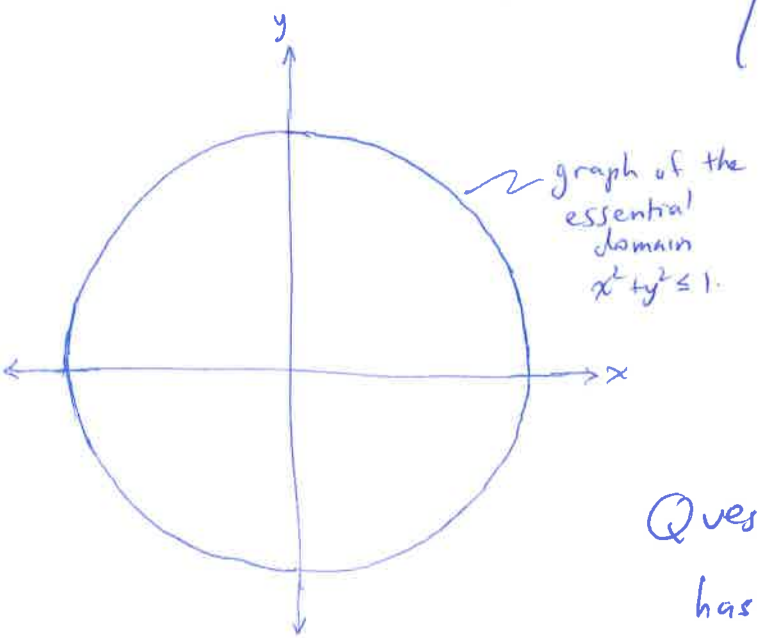
and similarly, the Conditional pdf of Y given $X=x$ is

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

which is only defined on the essential domain of f_X i.e. the region where $f_X(x) > 0$.

Ex. Suppose we throw a dart at a dartboard of radius $r=1$. and we let X be the x-coordinate of where the dart lands and Y be the y-coord. The one model for joint pdf is to assume each point on the dartboard is Equally-likely:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & \text{if } x^2+y^2 \le 1 \\ 0 & \text{if } x^2+y^2 > 1. \end{cases}$$



Question: If we are told the dart has x-coordinate $= 1/2$, what is the probability the y-coord. $\geq \frac{1}{2}$

i.e. Can we compute

$$P(Y \geq \tfrac{1}{2} \mid X = \tfrac{1}{2}).$$

Notice that $X=\frac{1}{2}$ is an event of probability 0 so we cannot say that $P(Y \geq \frac{1}{2} \mid X=\frac{1}{2}) = \dfrac{P(Y\geq\frac{1}{2}, X=\frac{1}{2})}{P(X=\frac{1}{2})}$ ← this is an indeterminant form $\frac{0}{0}$.

To compute this probability we need to find the conditional density of Y given $X = \frac{1}{2}$.

$$f_{Y|X}(y|\tfrac{1}{2}) = \frac{f_{X,Y}(\frac{1}{2}, y)}{f_X(\frac{1}{2})}.$$

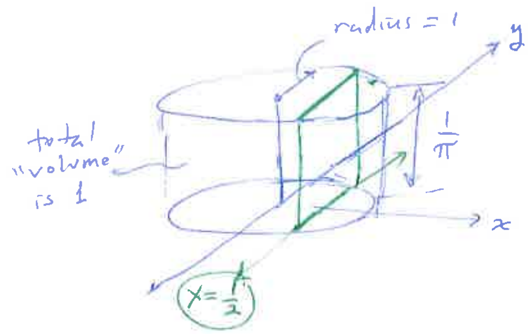
So we will need to find the marginal pdf of X:

for $-1 < x < 1$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dx = \frac{2\sqrt{1-x^2}}{\pi}.$$

and when $x = \frac{1}{2}$,

$$f_X(\tfrac{1}{2}) = \frac{2\sqrt{1-(\frac{1}{2})^2}}{\pi} = \frac{\sqrt{3}}{\pi}$$



Moreover,

$$f_{X,Y}(\tfrac{1}{2}, y) = \frac{1}{\pi} \quad \text{if} \quad -\frac{\sqrt{3}}{2} < y < \frac{\sqrt{3}}{2}.$$

So

$$f_{Y|X}(y|\tfrac{1}{2}) = \frac{\frac{1}{\pi}}{\frac{\sqrt{3}}{\pi}} = \frac{1}{\sqrt{3}} \quad \text{for} \quad -\frac{\sqrt{3}}{2} < y < \frac{\sqrt{3}}{2}$$

No surprise that Y conditionally has the uniform distribution

So

$$P(Y \geq \tfrac{1}{2} | X = \tfrac{1}{2}) = \int_{\frac{1}{2}}^{\frac{\sqrt{3}}{2}} \frac{1}{\sqrt{3}}\,dy = \left(\frac{\sqrt{3}}{2} - \frac{1}{2}\right)\frac{1}{\sqrt{3}} = \frac{\sqrt{3}-1}{2\sqrt{3}} \approx .211$$

$\blacksquare$

Remark In the last example, if we were given the dart landed in the 1st quadrant (i.e. where $X>0$, $Y>0$) what would the probability be of having $Y<\frac{1}{2}$?

i.e.

$$P(Y<\tfrac{1}{2} \mid X>0, Y>0).$$

Now, notice, $P(X>0, Y>0) = \frac{1}{4} > 0$ and we can use the conditional probability formula from earlier in the course:

$$P(Y<\tfrac{1}{2} \mid X>0, Y>0) = \frac{P(Y<\frac{1}{2}, X>0, Y>0)}{P(X>0, Y>0)}$$

$$= \frac{P(X>0,\ 0<Y<\frac{1}{2})}{P(X>0, Y>0)}$$

$$= \frac{\frac{1}{\pi}\int_0^{\frac{1}{2}} \sqrt{1-x^2}\,dx}{\frac{1}{4}} \underset{\substack{\text{trig. subst.} \\ x=\sin u \\ dx=\cos(u)\,du}}{=} \frac{\frac{1}{\pi}\int_0^{\frac{\pi}{6}} \cos^2(u)\,du}{\frac{1}{4}}$$

$$= \frac{\frac{1}{\pi}\int_0^{\frac{\pi}{6}} \frac{1+\cos(2u)}{2}\,du}{\frac{1}{4}} = \frac{4}{\pi}\left(\frac{\pi}{12}+\frac{\sqrt{3}}{8}\right) \approx .608998\ldots$$

Chapter 4 Derived distributions

Very often it is necessary to understand the distribution of a r.v. that is a function of some other r.v. whose distribution is known.

For example, suppose $U \sim \text{uniform}(0,1)$. What is the pdf of $Y = U^2$.

Here is one approach: it uses the fact that if we can find the cdf of Y, then its pdf can be gotten by differentiation:

Let's find (if possible/easy)

assume $y > 0$ since if $y \le 0$, $F_Y(y) = 0$.

$F_Y(y) = P(Y \le y)$ Note that since $U > 0$ then $Y > 0$, too

$= P(U^2 \le y)$

$= P(|U| \le \sqrt{y}) = P(U \le \sqrt{y})$ since $U > 0$. it is uniform(0,1)

$= \sqrt{y}$ since we know the cdf of a uniform(0,1) is $F_U(u) = u$ for $0 < u < 1$.

So $F_Y(y) = \sqrt{y}$ if $0 < y < 1$. $F_Y(y) = 0$ for $y \le 0$ and $F_Y(y) = 1$ for $y \ge 1$

Therefore, if $0 < y < 1$

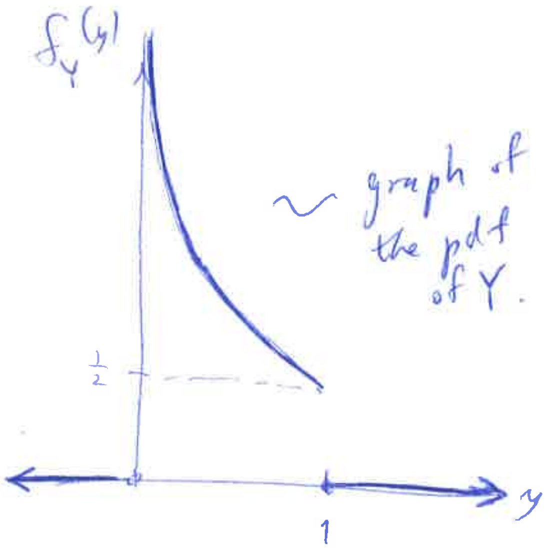
$$\frac{d}{dy} F_Y(y) = f_Y(y) = \frac{d}{dy}\left(\sqrt{y}\right) = \frac{1}{2\sqrt{y}}$$

and for $y \notin (0,1)$,

$$f_Y(y) = 0.$$

So the pdf of Y is

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & \text{for } 0 < y < 1 \\ 0 & \text{for other } y. \end{cases}$$



The CDF method for finding the pdf of a function of a continuous r.v (or r.vs.)

The basic problem is that we have a continuous r.v X (say) whose cdf we know (or can find) $F_X(x)$.
and we want to find the distribution (say, pdf) $f_Y(y)$ of some function of X; say $Y = g(X)$.

This can be completed the following way (However, this is not the only way, and this way may not be the easiest way).

1. Write down the cdf of $Y=g(X)$, i.e. $F_Y(y) = P(g(X) \leq y)$.

2. Undo (or Invert) the function g to "solve" for X and then take this expression and write it in terms of the "known" cdf of X.

3. Then take a derivative in y (using the chain rule) to obtain the pdf of Y $f_Y(y)$.

Here is an example when g is one-to-one (and therefore) has an inverse. We also suppose g is increasing.* ← If g is decreasing then we will handle this later.

$$F_Y(y) = P(Y \le y) = P(g(X) \le y)$$ ← substituted the information about Y, namely $Y = g(X)$

$$= P(X \le g^{-1}(y))$$ ← used the assumption that g^{-1} exists.

$$= F_X(g^{-1}(y))$$ ← this is just the cdf of X evaluated at $x = g^{-1}(y)$.

To get the pdf $f_Y(y)$ of Y we need to take a derivative in y:

$$\frac{d}{dy} F_Y(y) = f_Y(y) = \frac{d}{dy}\left(F_X(g^{-1}(y))\right)$$

$$= f_X(g^{-1}(y)) \frac{d}{dy}(g^{-1}(y))$$

Here's an application of this last result...

Suppose that $X \sim \text{Normal}(\mu, \sigma^2)$, i.e. X has the pdf

$$f(x) = \frac{e^{-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2}}{\sigma\sqrt{2\pi}} \quad \text{for } -\infty < x < \infty.$$

Let $Z = \dfrac{X-\mu}{\sigma} =: g(X)$.

Let's find the pdf of Z.

Notice that g is one-to-one, in fact,

$$g^{-1}(Z) = \sigma Z + \mu.$$

and $\dfrac{d}{dz} g^{-1}(z) = \sigma$

By the result on previous page

$$f_Z(z) = f_X(g^{-1}(z)) \frac{d}{dz}(g^{-1}(z))$$

$$= f_X(\sigma z + \mu) \cdot \sigma$$

$$= \frac{e^{-\frac{1}{2}\left\{\frac{(\sigma z+\mu)-\mu}{\sigma}\right\}^2}}{\sigma\sqrt{2\pi}} \cdot \sigma = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \quad \text{for } -\infty < z < \infty$$

That is, we recognize that Z has a Standard Normal distribution

We've Shown:

If $X \sim \text{Normal}(\mu, \sigma^2)$, then $Z = \dfrac{X-\mu}{\sigma} \sim \text{Normal}(0,1)$.

One more example ...

Suppose $Z \sim \text{Normal}(0,1)$ $\quad -\infty < z < \infty$

Find the pdf of $Y = Z^2$. — Note Z is not one-to-one on all of $\mathbb{R}$

for $y > 0$

$$F_Y(y) = P(Y \le y) = P(Z^2 \le y)$$

$$= P(|Z| \le \sqrt{y})$$

$$= P(-\sqrt{y} \le Z \le \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$

$$= 2\Phi(\sqrt{y}) - 1, \quad \text{where } \Phi(z) = \int_{-\infty}^{z} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\,du.$$

$$= 2\int_{-\infty}^{\sqrt{y}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\,du - 1$$

$$f_Y(y) = \frac{d}{dy}F_Y(y) = 2\left(\frac{e^{-\frac{(\sqrt{y})^2}{2}}}{\sqrt{2\pi}}\right)\frac{d}{dy}(\sqrt{y}) = \frac{\cancel{2}\, e^{-y/2}}{\sqrt{2\pi}} \cdot \frac{1}{\cancel{2}\sqrt{y}}$$

$$= \frac{y^{-\frac{1}{2}} e^{-y/2}}{2^{1/2}\sqrt{\pi}}$$

← this is the Chi-square distribution with 1 degree of freedom

i.e. Gamma$(\frac{1}{2}, 2)$.

Recall that if X and Y are <u>independent</u> jointly continuous r.v.s. then the joint pdf of X, Y factors as the product of its marginal pdfs:

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y). \quad \text{for all } x, y \in \mathbb{R}$$

and more generally if $X_1, X_2, \dots, X_n$ are <u>independent</u> jointly continuous then

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \quad \text{for all } x_1, \dots, x_n \in \mathbb{R}$$

Finding the distribution of the MAXIMUM of an independent collection of r.v.s.

We use the fact that

$$\left(\max\{X_1, X_2, \dots, X_n\} \le y \right) = \left(X_1 \le y, X_2 \le y, \dots, X_n \le y \right)$$

Then the cdf of $Y = \max\{X_1, \ldots, X_n\}$ is

$$F_Y(y) = P(Y \leq y)$$
$$= P(\max\{X_1, \ldots, X_n\} \leq y)$$
$$= P(X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y)$$ using fact on previous page

$$= P(X_1 \leq y)\, P(X_2 \leq y) \cdots P(X_n \leq y)$$ since the r.v.s are independent.

<u>Concrete example</u>. Suppose $X_1, X_2, \ldots, X_n$ are independent Uniform $(0,1)$ r.v.s. Then for $0 < y < 1$,

$$F_Y(y) = y^n \Rightarrow f_Y(y) = \begin{cases} n y^{n-1} & \text{for } 0 < y < 1 \\ 0 & \text{for other } y. \end{cases}$$

A similar approach can be applied when dealing with the $\min\{X_1, X_2, \ldots, X_n\}$. In this case we use the <u>fact</u> that

$$(\min\{X_1, X_2, \ldots X_n\} > y) = (X_1 > y, X_2 > y, \ldots, X_n > y)$$

So that if $W = \min(X_1, X_2, \ldots, X_n)$ where $X_1, X_2, \ldots, X_n$ are <u>independent</u>

$$F_W(w) = P(W \le w)$$

$$= 1 - P(W > w)$$

$$= 1 - P(\min\{X_1, X_2, \ldots, X_n\} > w)$$

$$= 1 - P(X_1 > w, X_2 > w, \ldots, X_n > w)$$

$$= 1 - P(X_1 > w)P(X_2 > w)\,P(X_3 > w) \cdots P(X_n > w)$$

<u>Concrete example</u> Suppose $X_1, X_2, \ldots, X_n \sim$ indep Uniform$(0,1)$

Then for $0 < w < 1$

$$F_W(w) = 1 - \{P(X_1 > w)\}^n = 1 - (1-w)^n$$

So

$$f_W(w) = \begin{cases} n\,(1-w)^{n-1} & \text{for } 0 < w < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Interesting application: Consider the r.v.

$$Y_n = n \cdot \min(X_1, X_2, \ldots, X_n)$$

Then if $X_1, \dots, X_n \sim$ indep. Uniform$(0,1)$. and $0<y<1$

$$F_{Y_n}(y) = P(n \cdot \min\{X_1, \dots, X_n\} \leq y)$$

$$= P\left(\min(X_1, \dots, X_n) \leq y/n\right)$$

$$= 1 - P\left(\min(X_1, \dots, X_n) > y/n\right)$$

$$= 1 - \left(1 - \frac{y}{n}\right)^n$$

and for $0<y<1$

$$f_{Y_n}(y) = n\left(1-\frac{y}{n}\right)^{n-1}\frac{1}{n} = \left(1-\frac{y}{n}\right)^{n-1} \text{ for } 0<y<1.$$

Notice that as $n \to \infty$

$$F_{Y_n}(y) \to 1 - e^{-y} \quad \text{as } n \to \infty$$

i.e the cdf of an exp(1). r.v.

$\blacksquare$

The CDF method extends to real-valued functions of more than 1 r.v.

Example

Suppose X and Y are independent exp(1) r.v.s

i.e. $$f(x,y) = \begin{cases} e^{-x} e^{-y} & \text{for } x>0, y>0 \\ 0 & \text{elsewhere} \end{cases}$$

Find the pdf of $W = X + Y$.

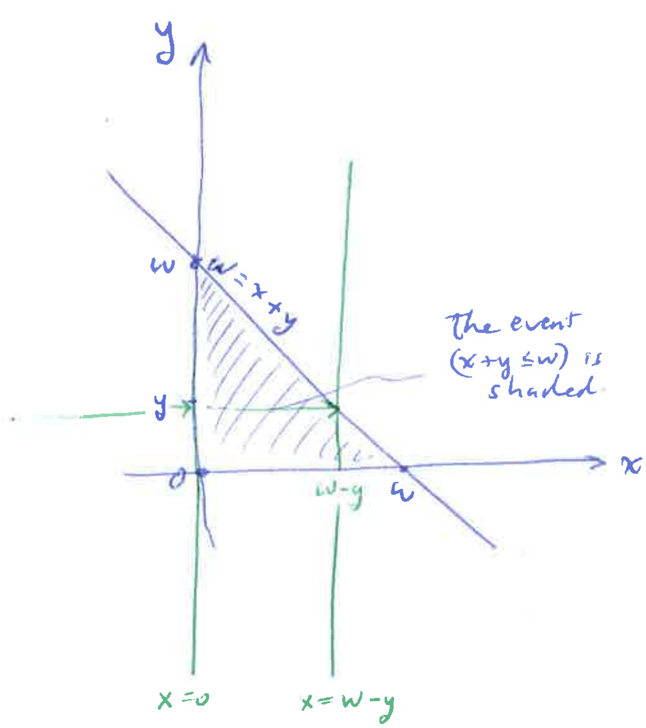
$$F_W(w) = P(W \le w) = P(X+Y \le w) = \int_0^w \int_0^{w-y} e^{-x} e^{-y}\, dx\, dy$$

$$= \int_0^w e^{-y} \left\{ \int_0^{w-y} e^{-x}\, dx \right\} dy$$

$$= \int_0^w e^{-y} \left\{ 1 - e^{-w+y} \right\} dy$$

$$= \int_0^w e^{-y}\, dy - \int_0^w e^{-w}\, dy$$

$$= 1 - e^{-w} - w e^{-w} \quad \text{for } w > 0.$$



$$f_W(w) = \frac{d}{dw} F_W(w)$$

$$= e^{-w} - e^{-w} + w e^{-w} = w e^{-w} \quad \text{for } w > 0.$$

$$= \frac{w^{2-1} e^{-w/1}}{1^2 \Gamma(2)} \sim \text{Gamma}(2,1).$$

Note: the exp(1) = Gamma(1,1). If we add two independent Gamma(1,1) distributions we get a Gamma(2,1) distribution.