Chapter 5 Notes: Distributions of Functions of Random Variables

Notes on Hogg-Tanis by Professor T. Fiore

(Next Page)

5.1 The Distribution Function Technique and Change of Variable Technique for Finding the Density or Mass Function of a Function of a Random Variable

In some situations, we have a random variable with a known p.d.f./p.m.f. and we would like to find the p.d.f./p.m.f. of a function of the random variable. For instance, if X is a standard normal, what is the density of $3X-1$?

We discuss now two methods:

1) The distribution function technique, which only works for continuous random variables
2) The change of variable technique, which works for continuous and discrete random variables, but with slightly different formulas.

In Section 5.4 we will also discuss

3) The moment generating function technique, which works well for linear combinations of random variables (all discrete or all continuous).

The Distribution Function Technique for Continuous Random Variables

Given a continuous random variable X (with support an interval and) with a known density, and another random variable Y expressed as a function $u(X)$, to find the density of Y do the following steps: 0) Find $\operatorname{supp} Y = u(\operatorname{supp} X)$.

1) In the c.d.f. $F_Y(y) = P(Y \leq y)$, rewrite $Y \leq y$ as $X \leq (\text{function of } y)$ using algebra and inverse functions.
2) Write this result in terms of an integral of the density of X for $y \in \operatorname{supp} Y$.
3) Differentiate the integral using the Fundamental Theorem of Calculus Part I and the Chain Rule, to obtain $F_Y'(y) = f_Y(y)$, the density of Y, for $y \in \operatorname{supp} Y$ and zero elsewhere.

<u>Ex</u> Let X be a standard normal random variable with density $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and consider $Y = 3X - 1$. Find the density of Y.

0) $u(x) = 3x - 1 \Rightarrow \operatorname{supp} Y = u(\operatorname{supp} X) = u((-\infty, \infty)) = (-\infty, \infty)$.
$= \mathbb{R}$

1)
$$F_Y(y) = P(Y \le y)$$
$$= P(3X - 1 \le y)$$
$$= P(3X \le y + 1)$$
$$= P\left(X \le \frac{y}{3} + \frac{1}{3}\right)$$

2)
$$= \int_{-\infty}^{\frac{y}{3}+\frac{1}{3}} f_X(x)\,dx \quad \text{for } y \in \overset{\text{supp } Y}{\overset{\|}{\mathbb{R}}}$$

$$= \int_{-\infty}^{\frac{y}{3}+\frac{1}{3}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx \quad \text{for } y \in \mathbb{R}$$

3)
$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{-\infty}^{\frac{y}{3}+\frac{1}{3}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx \quad \text{for } y \in \mathbb{R}$$
$$= \frac{1}{\sqrt{2\pi}} \cdot e^{-\left(\frac{y}{3}+\frac{1}{3}\right)^2/2} \cdot \frac{1}{3} \quad \text{for } y \in \mathbb{R}$$

Thus, the density of Y is

$$f_Y(y) = \frac{1}{3\sqrt{2\pi}} \cdot e^{-\left(\frac{y}{3}+\frac{1}{3}\right)^2/2} \quad \text{for } y \in \mathbb{R}$$

Ex Let X have the density $f_X(x) = \begin{cases} 2x & \text{if } x \in [0,1] \\ 0 & \text{else} \end{cases}$

and consider $Y = 3X - 1$. Find the density of Y using the distribution function technique.

(Notice: the formula for Y is the same as in the previous example, but this time the support of X is a finite length interval, rather than $\mathbb{R}$, so we must take some care with the integral bounds.)

$u(x) = 3x - 1$

The support of Y is $\operatorname{supp} Y = u(\operatorname{supp} X)$

$= u((0,1])$

b/c u is incr. $= (u(0), u(1)]$

$= (3\cdot 0 - 1, 3\cdot 1 - 1]$

$= (-1, 2]$

$$F_Y(y) = P(Y \le y)$$

$$= P(3X - 1 \le y)$$

$$= P(3X \le y + 1)$$

$$= P\left(X \le \tfrac{y}{3} + \tfrac{1}{3}\right)$$

$$(*) \quad = \int_{-\infty}^{\frac{y}{3}+\frac{1}{3}} f_X(x)\,dx$$

assume now $y \in \underset{\operatorname{supp} Y}{\overset{\shortparallel}{(-1,2]}}$ (this guarantees $\frac{y}{3} + \frac{1}{3} \in \operatorname{supp} X$)

$$= \int_0^{\frac{y}{3}+\frac{1}{3}} 2x\,dx \quad \text{for } y \in (-1,2]$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_0^{\frac{y}{3}+\frac{1}{3}} 2x\,dx = 2\left(\frac{y}{3} + \frac{1}{3}\right) \cdot \frac{1}{3} \quad y \in (-1,2]$$

$$= \frac{2}{9}y + \frac{2}{9} \quad y \in (-1,2]$$

Thus $f_Y(y) = \begin{cases} \frac{2}{9}y + \frac{2}{9} & \text{for } y \in [-1,2] \\ 0 & \text{otherwise} \end{cases}$

Rem In step (✱) we assumed $y \in (-1,2]$ and said that was enough to make sure that the integrand f_X would be non-zero there and only there. Let's double check to make sure that is the case.

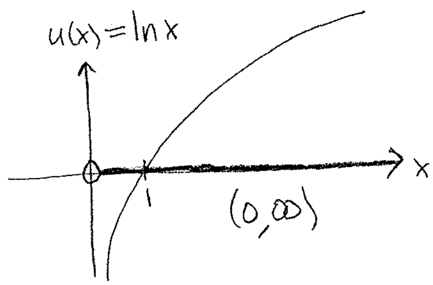
Simplify $0 < \frac{y}{3} + \frac{1}{3} \leq 1$ to have y by itself:

$$-\frac{1}{3} < \frac{y}{3} \leq \frac{2}{3}$$

$$-1 < y \leq 2 \quad \Rightarrow y \in (-1,2]$$, just as we found by computing $u((0,1])$.

Ex Let X be a random variable with an exponential distribution with mean θ. Suppose also that X does not have 0 in its image. Let $Y = \ln X$. Find the density of Y using the distribution function technique.

$\operatorname{supp} X = (0,\infty)$



$u((0,\infty)) = \mathbb{R}$
$\shortparallel$
$\operatorname{supp} Y$.

We have $f_X(x) = \begin{cases} 0 & \text{if } x \in (-\infty, 0] \\ \frac{1}{\theta} e^{-x/\theta} & \text{if } x \in (0, \infty) \end{cases}$

$$F_Y(y) = P(Y \le y) = P(\ln X \le y)$$

$$= P(e^{\ln X} \le e^y)$$

Recall: e^x is an increasing function, so "$a \le b \Rightarrow e^a \le e^b$"

$$= P(X \le e^y)$$

$$= \int_{-\infty}^{e^y} f_X(x)\, dx$$

$$= \int_0^{e^y} \frac{1}{\theta} e^{-x/\theta}\, dx \quad \text{assume } y \in \underset{\text{supp } Y}{\overset{\shortparallel}{\mathbb{R}}}$$

For $y \in \mathbb{R}$ we have

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_0^{e^y} \frac{1}{\theta} e^{-x/\theta}\, dx \quad \text{for } y \in \mathbb{R}$$

$$= \left(\frac{1}{\theta} e^{-e^y/\theta}\right) e^y \quad \text{for } y \in \mathbb{R}$$

Ex Let X be a random variable uniformly distributed on $[-1,2]$, i.e. X is $U(-1,2)$.

Consider $Y = X^2$. Find the density of Y.

$$f_X(x) = \begin{cases} \frac{1}{3} & \text{for } x \in [-1,2] \\ 0 & \text{otherwise} \end{cases}$$

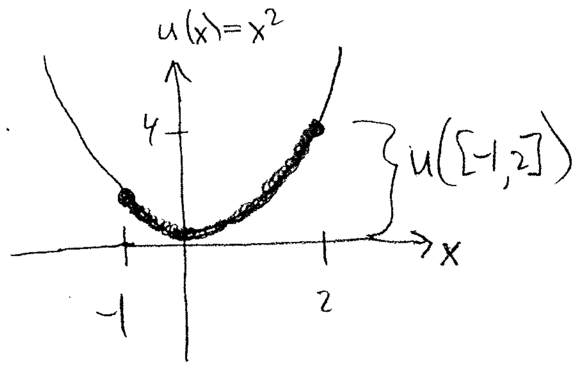
Notice: The image of X is $[-1,2]$ because the image of a random variable is the support of its density (or p.m.f in the discrete case). Then the image of $Y = X^2$ is the image of $[-1,2]$ under the function $u(x) := x^2$ (because $\text{image}(Y) = \text{image}(X^2)$

$$= \text{image}(u \circ X)$$

$$= u(\text{image}(X))$$

$$= u([-1,2])$$

$$= [0,4] \quad)$$



Thus, the image of Y is $[0,4]$ and the density f_Y will be 0 outside of $[0,4]$.

$$F_Y(y) = P(Y \le y)$$

$$= P(X^2 \le y)$$ this is zero if $y<0$ because $X^2 \ge 0$, so assume now $y \ge 0$.

$$= P(\sqrt{X^2} \le \sqrt{y})$$ because $\sqrt{\ }: \mathbb{R}_{\ge 0} \to \mathbb{R}$ is an increasing function

$$= P(|X| \le \sqrt{y})$$ because $\sqrt{X^2} = |X|$.

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

$$= \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx$$ (we make no assumption on y this time!)

$$\Rightarrow \quad f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx$$

$$= \frac{d}{dy}\left(\int_{-\sqrt{y}}^{0} f_X(x)\,dx + \int_{0}^{\sqrt{y}} f_X(x)\,dx \right)$$

$$f_Y(y) = \frac{d}{dy}\left(-\int_0^{-\sqrt{y}} f_X(x)\,dx + \int_0^{\sqrt{y}} f_X(x)\,dx\right)$$

$$f_Y(y) = +f_X(-\sqrt{y})\cdot\frac{-1}{2\sqrt{y}} + f_X(\sqrt{y})\cdot\frac{1}{2\sqrt{y}}$$

by FTC1 and Chain Rule.

Recall that f_X is piecewise defined in this example (X is unif. on $[-1,2]$) and that f_Y has support $[0,4]$, so we consider that in computing $f_X(-\sqrt{y})$ and $f_X(\sqrt{y})$ for $y \geq 0$.

For $\underbrace{-1 \leq -\sqrt{y} \leq 0}$ we have $f(-\sqrt{y}) = \frac{1}{3}$, otherwise zero.

$(-1)^2 \geq (-\sqrt{y})^2 \geq 0$ (x^2 is decr. on $(-\infty, 0]$).

$1 \geq y \geq 0.$

For $0 \leq y \leq 4$, $f_X(\sqrt{y}) = \frac{1}{3}$, otherwise zero.

From above,

$$f_Y(y) = +\frac{1}{2\sqrt{y}} f(-\sqrt{y}) + \frac{1}{2\sqrt{y}} f(\sqrt{y}) \qquad \text{for } y \geq 0$$

$$f_Y(y) = \begin{cases} +\frac{1}{2\sqrt{y}} \cdot \frac{1}{3} + \frac{1}{2\sqrt{y}} \cdot \frac{1}{3} & \text{for } 0 \leq y \leq 1 \\ 0 + \frac{1}{2\sqrt{y}} \cdot \frac{1}{3} & \text{for } 1 \leq y \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

This concludes our treatment of the distribution function technique for finding the density of a function of a given random variable.

Next we see another technique: the change of variable technique, which has formulas for continuous and discrete cases.

The Change of Variable Technique for Continuous Random Variables

Thm Let X be a continuous random variable with density f_X and support (c_1, c_2) (c_1 could be $-\infty$ and c_2 could be ∞). Suppose $u:(c_1,c_2)\to\mathbb{R}$ is differentiable and either strictly increasing or strictly decreasing on (c_1,c_2), the support of X. Let $v: u((c_1,c_2))\to(c_1,c_2)$ be the inverse of u. Then the density of $Y = u\circ X$ is

$$f_Y(y) = \begin{cases} f_X(v(y))\,|v'(y)| & \text{for } y\in u((c_1,c_2)) \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the support of Y is the u-image of (c_1,c_2).

Rem If $u:(c_1,c_2)\to\mathbb{R}$ is increasing, then the image of u is $(u(c_1), u(c_2))$, and this is the image of Y. If $u:(c_1,c_2)\to\mathbb{R}$ is decreasing, then the image of u is $(u(c_2), u(c_1))$, and this is the image of Y. ("$u(c_i)$" may need to be interpreted as $\lim_{x\to c_i} u(x)$, Ex $\ln(0,\infty)=(-\infty,\infty)$)

Ex Let X be a standard normal random variable with density $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and consider $Y = 3X - 1$. Find the density of Y using the change of variable technique.

~~$(c_1, c_2) = (-\infty, \infty)$, $u(x) = 3x - 1$ is strictly increasing~~ $(u'(x) = 3 > 0 \ \forall x \in \operatorname{supp} X)$.

$$y = 3x - 1$$
$$y + 1 = 3x$$
$$\frac{y}{3} + \frac{1}{3} = x$$

The support of Y is $u((-\infty,\infty)) = (-\infty, \infty) = \mathbb{R}$

$\Rightarrow$ $v(y) = \frac{y}{3} + \frac{1}{3}$ is inverse to u.

Thm $\Rightarrow$

$$\begin{aligned}
f_Y(y) &= f_X(v(y)) \cdot |v'(y)| && \text{for } y \in \mathbb{R} = \operatorname{supp} Y \\
&= f_X\left(\frac{y}{3} + \frac{1}{3}\right) \cdot \left|\frac{1}{3}\right| && \text{for } y \in \mathbb{R} \\
&= \frac{1}{\sqrt{2\pi}} \cdot e^{-\left(\frac{y}{3} + \frac{1}{3}\right)^2/2} \cdot \frac{1}{3} && \text{for } y \in \mathbb{R} \\
&= \frac{1}{3\sqrt{2\pi}} \cdot e^{-\left(\frac{y}{3} + \frac{1}{3}\right)^2/2} && \text{for } y \in \mathbb{R}
\end{aligned}$$

$\boxed{\text{same answer as distribution function technique!}}$

Ex Let X have the density $f_X(x) = \begin{cases} 2x & \text{if } x \in (0,1) \\ 0 & \text{else} \end{cases}$
and consider $Y = 3X - 1$. Find the density of Y using the change of variable technique.

$(c_1, c_2) = (0,1)$, $u(x) = 3x - 1$ is strictly increasing
$(u'(x) = 3 > 0 \ \forall x \in (0,1))$

$y = 3x - 1$
$y + 1 = 3x$
$\frac{y}{3} + \frac{1}{3} = x$

The support of Y is $u((0,1)) = (u(0), u(1))$
$= (-1, 2)$

$\Rightarrow v(y) = \frac{y}{3} + \frac{1}{3}$ is inverse to u.

$$\text{Thm} \Rightarrow f_Y(y) = f_X(v(y)) \cdot |v'(y)| \quad \text{for } y \in (-1,2)$$
$$= f_X\left(\frac{y}{3} + \frac{1}{3}\right) \cdot \frac{1}{3} \quad \text{for } y \in (-1,2)$$
$$= 2\left(\frac{y}{3} + \frac{1}{3}\right) \cdot \frac{1}{3} \quad \text{for } y \in (-1,2)$$
$$= \frac{2}{9}y + \frac{2}{9} \quad \text{for } y \in (-1,2)$$

(zero otherwise)

The support of Y is the u-image of $(0,1)$:
$$u((0,1)) \underset{\substack{u \text{ is increasing} \\ \text{and defined at} \\ \text{endpoints.}}}{=} (u(0), u(1)) = (3\cdot 0 - 1, 3\cdot 1 - 1) = (-1, 2)$$

Thus, the density of Y is $f_Y(y) = \begin{cases} \frac{2}{9}y + \frac{2}{9} & \text{for } y \in (-1,2) \\ 0 & \text{otherwise} \end{cases}$

Ex Let X be a random variable with an exponential distribution with mean θ. Suppose also that X does not have 0 in its image. Let $Y = \ln X$. Find the density of Y using the change of variable technique.

$$f_X(x) = \begin{cases} 0 & \text{if } x \in (-\infty, 0] \\ \frac{1}{\theta} e^{-x/\theta} & \text{if } x \in (0, \infty) \end{cases}.$$

$(c_1, c_2) = (0, \infty)$

$u(x) = \ln x$

$u'(x) = \frac{1}{x} > 0$ for $x \in (0, \infty)$

$\Rightarrow u$ is increasing on $\operatorname{supp}(X)$.

$\operatorname{supp} Y = u((0,\infty)) = \mathbb{R}$ (see next page or recall earlier example).

$y = \ln x$

$e^y = e^{\ln x}$

$e^y = x \quad \Rightarrow \quad v(y) = e^y$ is the inverse to $u(x)$.

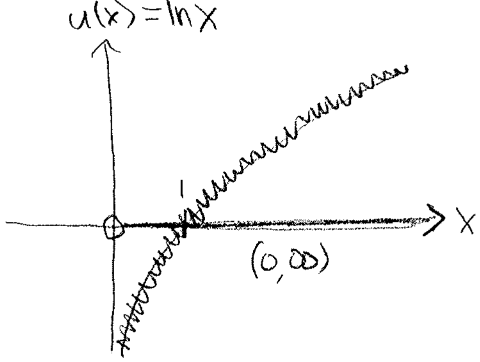
$v'(y) = e^y$

$$\text{Thm} \Rightarrow f_Y(y) = f_X(v(y)) \cdot |v'(y)| \quad \text{for } y \in \mathbb{R}$$

$$= f_X(e^y)\, e^y \quad \text{for } y \in \mathbb{R}$$

$$= \left(\frac{1}{\theta} e^{-e^y}\right) e^y \quad \text{for } y \in \mathbb{R}$$

But why is $\mathrm{supp}(Y) = u((0,\infty)) \overset{?}{=} \mathbb{R}$?



$$u((0,\infty)) = \mathbb{R} = \mathrm{supp}\, Y$$

Thus $f_Y(y) = \left(\frac{1}{\theta} e^{-e^y}\right) e^y$ for all $y \in \mathbb{R}$.

<u>Rem</u> In our version of the change of variable technique for continuous random variables, we assumed u was either strictly increasing or strictly decreasing on the support of the random variable X. So for instance, the theorem would not apply to $u(x) = x^2$ when $\mathrm{supp}(X) = (-1, 2)$, because u is first decreasing on $(-1, 0)$ and then increasing on $(0, 2)$. Nevertheless, a generalization of the theorem can be applied in such a situation. We don't go into it, but if you're curious, see page 222 of the textbook.

The Change of Variable Technique for Discrete Random Variables

Thm Let X be a discrete random variable with probability mass function f_X.

Let $u: \text{supp}(X) \longrightarrow \mathbb{R}$ be an injective function.

Let $v: u(\text{supp}(X)) \longrightarrow \text{supp}(X)$ be the inverse to u.

Then the pmf of $Y = u \circ X$ is

$$f_Y(y) = f_X(v(y)) \quad \text{for } y \in u(\text{supp}(X)).$$

In particular, the support of Y is the u-image of the support of X.

Ex Let X be a discrete random variable with p.m.f.

x	1	2	3	4
$f_X(x)$	$1/10$	$2/10$	$3/10$	$4/10$

and consider $Y = X^2$.

Find the p.m.f. of Y.

5.2 The Change of Variables Technique for Finding the Joint Density of a Transformation of Two Random Variables

Sometimes one has two jointly distributed random variables X and Y with known joint density and one would like to find the joint density of $u(X,Y)$ and $v(X,Y)$. For instance we may know the joint density of X and Y and wish to find the joint density of $U := X - Y$ and $V := X + Y$. In this section we learn a change of variable technique for just this situation.

Warning! The approach in these notes does not follow the book. Instead of multiplying $f_{X,Y}$ by the absolute value of the Jacobian of the inverse transformation like the book, we divide $f_{X,Y}$ by the absolute value of the Jacobian of original transformation. Also, instead of X_1, X_2 and Y_1, Y_2 as in the book, we use X, Y and U, V.

Reason Our method is easier to use in practice.

We begin with a rapid review of the background on joint densities of jointly distributed continuous random variables, from section 4.1.

Review of Joint Densities for Continuous Random Variables

Two random variables are called jointly distributed if they are defined on the same sample space. If X and Y are jointly distributed, then we have a function into $\mathbb{R}^2$, namely

$$(X,Y): S \longrightarrow \mathbb{R}^2$$
$$s \longmapsto (X(s), Y(s)).$$

The image of this function is called the support of X and Y. We denote this image by $\operatorname{supp}(X,Y)$. This image is not necessarily the product of $\operatorname{supp}(X)$ and $\operatorname{supp}(Y)$, see for example Example 4.1-1 on page 180 and Example 5.2-1 on pages 225-226.

There is a 2-dimensional version of the notion of density.

A joint probability density function (if one exists) for jointly distributed continuous random variables X and Y is an integrable function $f_{X,Y}: \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that:

1) $f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$,

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

3) If D is a reasonable subset of $\mathbb{R}^2$, then $P((X,Y) \in D) = \iint_D f_{X,Y}\, dA$,

in other words the probability that (X,Y) is in the region D is the volume of the solid above the region D and below the surface $z = f_{X,Y}(x,y)$,

4) $f_{X,Y}(x,y) \neq 0$ if and only if the point (x,y) is in $\mathrm{supp}(X,Y)$.

Change of Variable Technique for a Transformation of Two Continuous Random Variables

Thm Let X and Y be jointly distributed continuous random variables with joint density $f_{X,Y}$, and assume $\text{supp}(X,Y)$ is open.

Let
$$T : \text{supp}(X,Y) \to \mathbb{R}^2$$
$$(x,y) \longmapsto (u(x,y), v(x,y))$$

be an injective function such that

1. T_1 and T_2 have continuous partial derivatives at all points in $\text{supp}(X,Y)$,

2. The determinant
$$\text{Jac}(T) := \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

is non-zero at all points in $\text{supp}(X,Y)$.

Let $U := u(X,Y)$ and $V := v(X,Y)$.

Then, $\text{supp}(U,V) = T(\text{supp}(X,Y))$

and the joint density for U and V is

$$f_{U,V}(u,v) = f_{X,Y}(x(u,v), y(u,v)) / |\mathrm{Jac}(T)|$$

for $(u,v) \in \mathrm{supp}(U,V)$.

where, after simplifying in x and y on the right hand side, we replace x and y by $x(u,v)$ and $y(u,v)$ from $(x(u,v), y(u,v)) = T^{-1}(u,v)$.

<u>Ex</u> Let X and Y be jointly distributed exponential random variables with mean 1, so that

$$f_X(x) = e^{-x} \text{ for } x \in (0,\infty)$$

and $f_Y(y) = e^{-y}$ for $y \in (0,\infty)$.

Suppose X and Y are independent, which means

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

and $\mathrm{supp}(X,Y) = \mathrm{supp}(X) \times \mathrm{supp}(Y)$

$$= \{(x,y) \in \mathbb{R}^2 \mid x \in \mathrm{supp}\, X \text{ and } y \in \mathrm{supp}\, Y\}.$$

Let $U = X - Y$ and $V = X + Y$. <u>Find</u> $f_{U,V}$.

$$T(x,y) := (x-y,\ x+y) = (u(x,y),\ v(x,y)),$$

$T: (0,\infty)\times(0,\infty) \to \mathbb{R}$ is injective.

$$\mathrm{Jac}(T) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 1-(-1) = 2 \neq 0$$

In $u = x-y$, $v = x+y$, solve for x and y in terms of u, v to determine T^{-1}!

$u + y = x \Rightarrow v = (u+y)+y$

$v = u + 2y$

$\boxed{\frac{v-u}{2} = y}$ Plug into $u+y = x$

$u + \left(\frac{v-u}{2}\right) = x$

$\boxed{\frac{u+v}{2} = x}$

Thm $\Rightarrow$

$$f_{U,V}(u,v) = f_{X,Y}(x,y) / |\mathrm{Jac}(T)| \quad \text{for } (u,v)\in \mathrm{Supp}(U,V)$$

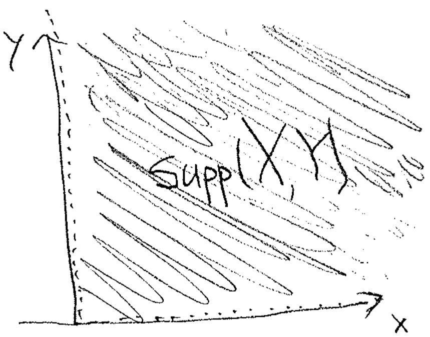
$$= e^{-x}\cdot e^{-y}/|2| \quad \text{for } (u,v)\in \mathrm{supp}(U,V)$$

$$= \frac{1}{2}e^{-(x+y)} = \frac{1}{2}e^{-\left(\frac{u+v}{2}+\frac{v-u}{2}\right)} = \boxed{\frac{1}{2}e^{-v}}$$

for $(u,v)\in \mathrm{supp}(U,V)$.

But how do we find $\operatorname{supp}(U,V)$?

By the Thm, $\operatorname{supp}(U,V) = T(\operatorname{supp}(X,Y))$, so to find $\operatorname{supp}(U,V)$, we need to find the image of

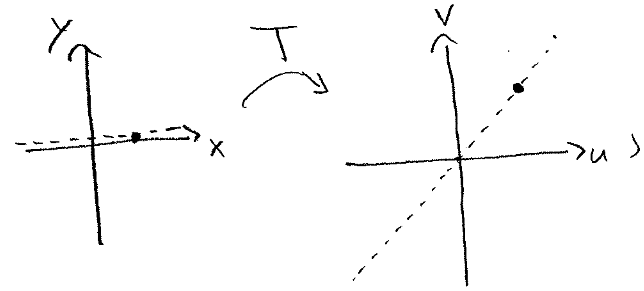


under the transformation T.

Look at what T does to the boundaries:

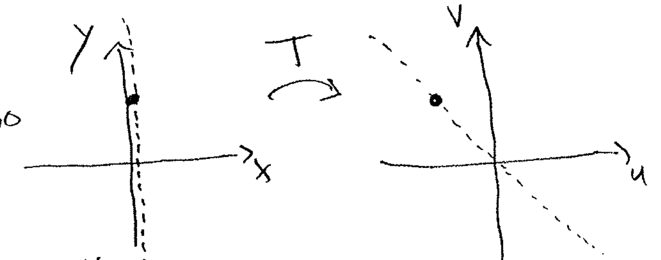
$T(x,y) = (x-y,\, x+y)$ so

$T(x,0) = (x-0,\, x+0) = (x,x)$ so

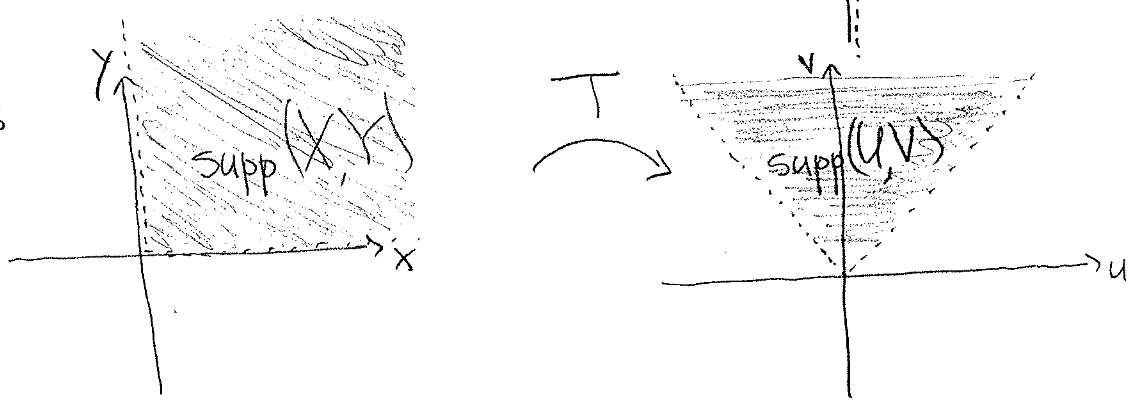


similarly

$T(0,y) = (0-y,\, 0+y) = (-y,y)$ so



Thus



$\operatorname{supp}(U,V) = \{(u,v) \in \mathbb{R}^2 \mid v \in (0,\infty) \text{ and } u \in (-v,v)\}$. $\square$

Rem For comparison to how the book would do it, see Example 5.2-2 on pages 226-227.

Ex Let X and Y be independent and identically distributed exponential random variables with parameter $\lambda = 4$, i.e. $\theta = 1/4$.

Then $f_{X,Y}(x,y) = f_X(x) f_Y(y)$

$$= 4e^{-4x} \cdot 4e^{-4y} \quad \text{for } x,y>0.$$

$$= 16e^{-4(x+y)} \quad \text{for } x,y>0.$$

(a) Let $U = 2X+Y$, $V = X/Y$.

Find $f_{U,V}(u,v)$.

$$u(x,y) = 2x+y, \quad v(x,y) = x/y$$

$$\mathrm{Jac}(T) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ \frac{1}{y} & -\frac{x}{y^2} \end{vmatrix} = \frac{-2x-y}{y^2}$$

In $u = 2x + y$, $v = x/y$, solve for x and y in terms of u, v to determine T^{-1}.

$$u - 2x = y \implies v = x/y$$
$$\Downarrow$$
$$v = x/(u-2x)$$
$$uv - 2xv = x$$
$$uv - 2xv - x = 0$$
$$uv - x(2v+1) = 0$$
$$uv = x(2v+1)$$
$$\boxed{\frac{uv}{2v+1} = x} \xrightarrow{\text{Plug into}} u - 2x = y$$
$$u - 2\left(\frac{uv}{2v+1}\right) = y$$
$$\frac{2uv+u}{2v+1} - \frac{2uv}{2v+1} = y$$
$$\boxed{\frac{u}{2v+1} = y}$$

$$\text{Thm} \implies f_{U,V}(u,v) = f_{X,Y}(x,y) \Big/ |\mathrm{Jac}(T)| \quad \text{for } (u,v) \in \mathrm{supp}(U,V)$$
$$= 16e^{-4(x+y)} \Big/ \left| \frac{-2x-y}{y^2} \right| \quad \text{for } (u,v) \in \mathrm{supp}(U,V)$$

algebra!

$$= 16e^{-4\left(\frac{u(v+1)}{1+2v}\right)} \cdot \left| \frac{-u}{(2v+1)^2} \right| \quad \text{for } (u,v) \in \mathrm{supp}(U,V)$$

$$\boxed{\begin{array}{l} x, y > 0 \\ \implies u = 2x + y > 0 \end{array}}$$

$$= 16e^{-4\left(\frac{u(v+1)}{1+2v}\right)} \cdot \left(\frac{u}{(2v+1)^2} \right) \quad \text{for } (u,v) \in \mathrm{supp}(U,V).$$

□

Clever Application of Change of Variables Technique to Find the Density of a Function of Two Random Variables

Sometimes one has two random variables with known densities and would like to find the density of a single function of the two given random variables. This can be found using the change of variable technique with U = function, V = something appropriate, and then finding the marginal density f_U with $f_{U,V}$.

<u>Ex</u> (F distribution) 5.2 #2 Let X and Y be independent chi-square random variables with q and r degrees of freedom, respectively.

$$f_X(x) = \frac{1}{\Gamma(q/2)\, 2^{q/2}} x^{\frac{q}{2}-1} e^{-x/2}, \quad f_Y(y) = \frac{1}{\Gamma(r/2)\, 2^{r/2}} y^{\frac{r}{2}-1} e^{-y/2}$$

The random variable $\dfrac{X/q}{Y/r}$ is said to have an <u>F-distribution with q and r degrees of freedom</u> (the numerator degrees of freedom always comes first in the name). The F distribution is used in the analysis of variance (ANOVA).

We now find the density of $\dfrac{X/q}{Y/r}$ using the change of variable technique.

Let $U = \dfrac{X/q}{Y/r}$ and $V = Y$.

$$u(x,y) = \frac{x/q}{y/r} \qquad v(x,y) = y$$

$$\mathrm{Jac}(T) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{1/q}{y/r} & \frac{-x/q}{y^2/r} \\ 0 & 1 \end{vmatrix}$$

$$= \frac{1/q}{y/r} = \frac{r}{yq}$$

In $u = \frac{x/q}{y/r}$, $v = y$ solve for x and y in terms of u, v to determine T^{-1}.

$$u = \frac{x/q}{y/r} \qquad \boxed{v = y}$$

$$\Rightarrow \quad \Downarrow$$

$$u = \frac{x/q}{v/r} \qquad \text{supp}(U,V) = (0,\infty)\times(0,\infty)$$

$$\frac{uv}{r} = \frac{x}{q}$$

$$\boxed{\frac{uvq}{r} = x}$$

$$\text{Thm} \Rightarrow f_{U,V}(u,v) = f_{X,Y}(x,y) \Big/ |\text{Jac}(T)| \quad (u,v)\in\text{supp}(U,V)$$

$$= f_X(x) f_Y(y) \Big/ \left|\frac{r}{yq}\right| \quad \text{for } (u,v)\in\text{supp}(U,V)$$

$$= c\, x^{\frac{q}{2}-1} \cdot y^{\frac{r}{2}-1} \cdot e^{-(x+y)/2} \cdot \frac{yq}{r} \quad \text{for } (u,v)\in\text{supp}(U,V)$$

where $c = \dfrac{1}{\Gamma(q/2)\Gamma(r/2)2^{\frac{q+r}{2}}}$

$$= c\left(\frac{uvq}{r}\right)^{\frac{q}{2}-1}\cdot\left(v\right)^{\frac{r}{2}-1}\cdot e^{-\left(\frac{uvq}{r}+v\right)/2}\cdot\frac{vq}{r} \quad \text{for } (u,v)\in\operatorname{supp}(U,V)$$

$$f_{U,V}(u,v) = c\left(\frac{q}{r}\right)^{\frac{q+r}{2}}\cdot u^{\frac{q}{2}-1}\cdot\left(v^{\frac{q+r}{2}-1}\cdot e^{-\left(1+\frac{uq}{r}\right)v/2}\right) \quad \text{for } (u,v)\in\operatorname{supp}(U,V)$$

Then we find the marginal density f_U to find the F-density:

$$f_U(u) = \int_0^\infty f_{U,V}(u,v)\,dv \quad \text{for } u\in(0,\infty)$$

$$= c\left(\frac{q}{r}\right)^{\frac{q+r}{2}}\cdot u^{\frac{q}{2}-1}\int_0^\infty v^{\frac{q+r}{2}-1}\cdot e^{-\left(1+\frac{uq}{r}\right)v/2}\,dv \quad \text{for } u\in(0,\infty)$$

Substitute $w=\left(1+\frac{uq}{r}\right)v$

$dw=\left(1+\frac{uq}{r}\right)dv$

$$= c\left(\frac{q}{r}\right)^{\frac{q+r}{2}}\cdot u^{\frac{q}{2}-1}\int_0^\infty\left(\frac{w}{\left(1+\frac{uq}{r}\right)}\right)^{\frac{q+r}{2}-1}\cdot e^{-w/2}\cdot\frac{1}{\left(1+\frac{uq}{r}\right)}\,dw \quad u\in(0,\infty)$$

$$= c\left(\frac{q}{r}\right)^{\frac{q+r}{2}}\cdot u^{\frac{q}{2}-1}\cdot\frac{1}{\left(1+\frac{uq}{r}\right)^{\frac{q+r}{2}}}\cdot\underbrace{\int_0^\infty w^{\frac{q+r}{2}}\cdot e^{-w/2}\,dw}_{\Gamma\left(\frac{q+r}{2}\right)2^{\frac{q+r}{2}}\cdot 1} \quad u\in(0,\infty)$$

by comparison with a $\chi^2(q+r)$ density

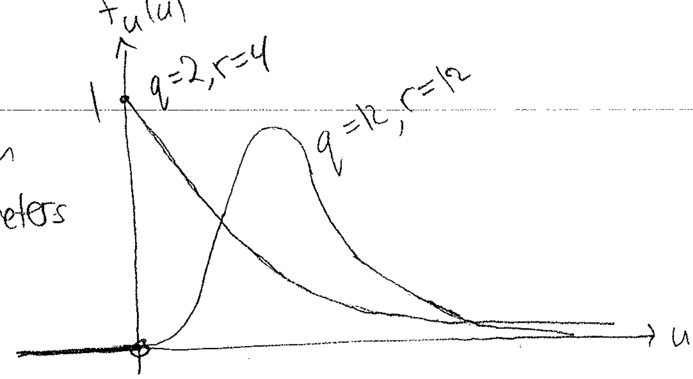
Next replace also c, and cancel $2^{\frac{q+r}{2}}$.

Thus, the F-density is

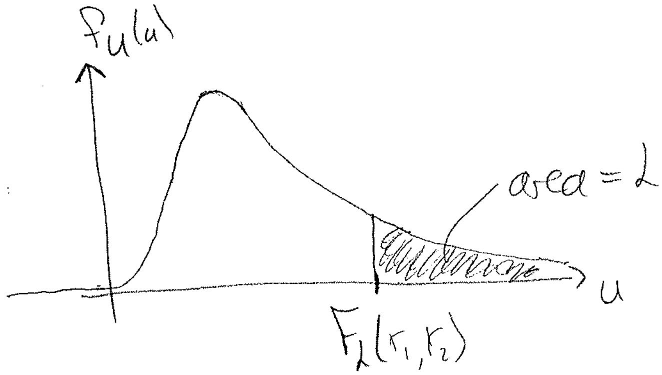
$$f_U(u) = \left(\frac{q}{r}\right)^{\frac{q+r}{2}} \cdot \left(\frac{\Gamma\left(\frac{q+r}{2}\right)}{\Gamma\left(\frac{q}{2}\right)\cdot\Gamma\left(\frac{r}{2}\right)}\right) \cdot \frac{u^{\frac{q}{2}-1}}{\left(1+\frac{uq}{r}\right)^{\frac{q+r}{2}}}$$

for $u \in (0,\infty)$ and zero otherwise.

Two sketches of F-densities with the specified parameters are:



Notation If a random variable has an F-distribution with r_1 and r_2 degrees of freedom, then we say its distribution is $F(r_1, r_2)$. For $0 \leq \alpha \leq 1$, the number $F_\alpha(r_1, r_2)$ is the input for which the right tail probability is α.



R: density is df(x, df1 = r_1, df2 = r_2)
c.d.f. is pf(q, df1 = r_1, df2 = r_2)
quantile is qf(p, df1 = r_1, df2 = r_2)

and $F_\alpha(r_1, r_2)$ is

qf(α, df1=r_1, df2=r_2, lower.tail=FALSE).

Ex 5.2 #4 Suppose the random variable W has a $F(9,24)$ distribution. Then

$F_{.05}(9,24)$ is

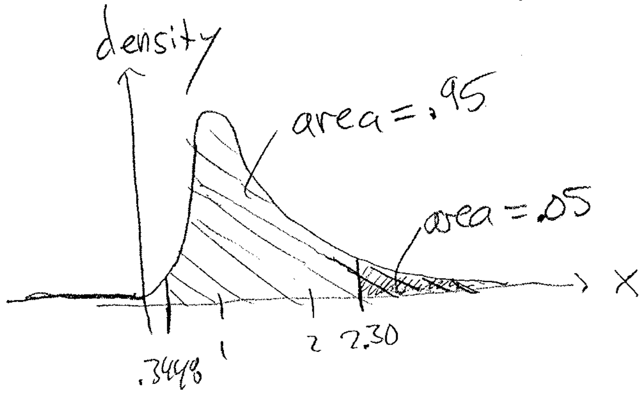
qf(.05, df1=9, df2=24, lower.tail=FALSE)

which gives 2.30

$F_{.95}(9,24)$ is

qf(.95, df1=9, df2=24, lower.tail=FALSE)

which gives .3448



To plot the pdf, type in

curve(df(x, df1=9, df2=24), -1, 4, lwd=2, col="red")

$P(.277 \le W \le 2.70)$ is

pf(2.7, df1=9, df2=24) - pf(.277, df1=9, df2=24)

Lemma $F_{1-\alpha}(r_1, r_2) = \dfrac{1}{F_{\alpha}(r_1, r_2)}$

Rem We have derived the F-density via a clever application of the change of variable technique, and then learned how to compute related quantities in R. Problem 5.2 #14 is another instance of this clever application: X and Y have given densities, then we find $U = X/Y$ using a change of variables with $U = X/Y$ and $V = Y$.

5.3 Mean and Variance of a Function of Several Independent Random Variables

We next learn how to find the mean and variance of linear combi's and products of independent random variables.

<u>Thm</u> Suppose $X_1, X_2, \ldots, X_n$ are random variables, and $a_1, a_2, \ldots, a_n$ are real constants. Then

(1) $$E(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n).$$

(<u>not</u> <u>assuming</u> <u>independence</u>!)

(2) Suppose now that $X_1, X_2, \ldots, X_n$ are <u>independent</u> random variables. Then

$$Var(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + \cdots + a_n^2 Var(X_n).$$

<u>Ex</u> Consider the roll of two 3-sided dice, one red, one blue. Let X_1 be the number on the red die and X_2 the number on the blue die. Let $Y := X_1 + X_2$. Compute $E(Y)$ and $Var(Y)$

in two ways: First directly from the p.m.f. for Y, then using the Theorem.

1) Find the p.m.f. for Y, $f_Y : \{2,3,4,5,6\} \to [0,1]$.

$$f_Y(2) = P(X_1 = 1 \text{ and } X_2 = 1) = P(X_1 = 1)\,P(X_2 = 1)$$
$$= \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

$$f_Y(3) = P(X_1 = 1, X_2 = 2 \text{ or } X_1 = 2, X_2 = 1)$$
$$= \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{2}{9}$$

$$f_Y(4) = P(X_1 = 1, X_2 = 3 \text{ or } X_1 = 2, X_2 = 2 \text{ or } X_1 = 3, X_2 = 1)$$
$$= \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{3}{9}$$

$$f_Y(5) = P(X_1 = 2, X_2 = 3 \text{ or } X_1 = 3, X_2 = 2)$$
$$= \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{2}{9}$$

$$f_Y(6) = P(X_1 = 3, X_2 = 3) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.$$

$$E(Y) = \sum_{y=2}^{6} y\, f_Y(y) = 2 \cdot \frac{1}{9} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{3}{9} + 5 \cdot \frac{2}{9} + 6 \cdot \frac{1}{9} = \frac{36}{9} = \boxed{4}$$

$$\mathrm{Var}(Y) = \sum_{y=2}^{6} (y-4)^2 f_Y(y) = 4 \cdot \frac{1}{9} + 1 \cdot \frac{2}{9} + 0 \cdot \frac{3}{9} + 1 \cdot \frac{2}{9} + 4 \cdot \frac{1}{9} = \frac{12}{9} = \boxed{\frac{4}{3}}$$

2) Find $E(X_1)$, $E(X_2)$, $Var(X_1)$, $Var(X_2)$ and use the Theorem.

$$E(X_1) = \sum_{x=1}^{3} x\, f_{X_1}(x) = 1\cdot\frac{1}{3} + 2\cdot\frac{1}{3} + 3\cdot\frac{1}{3} = 2$$

$$Var(X_1) = \sum_{x=1}^{3} (x-2)^2 f_{X_1}(x) = 1\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3} = \frac{2}{3}$$

Notice the same computations hold for X_2.

$$\text{Thm} \Rightarrow E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2)$$
$$= 2 + 2$$
$$= 4$$

same answer!

$$\text{Thm} \Rightarrow Var(Y) = Var(X_1 + X_2)$$
$$= 1^2 Var(X_1) + 1^2 Var(X_2)$$
$$= \frac{2}{3} + \frac{2}{3}$$
$$= \frac{4}{3}$$

same answer!

Next we have a formula for the expectation of of a function of independent random variables.

<u>Thm</u> Let $X_1, X_2, \dots, X_n$ be <u>independent</u> random variables, and $u(X_1, X_2, \dots, X_n)$ a function of $X_1, X_2, \dots, X_n$.

1) If they are discrete with respective p.m.f.'s $f_i(x_i)$, then

$$E(u(X_1, \dots, X_n)) = \sum_{x_1} \cdots \sum_{x_n} u(x_1, \dots, x_n) f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

2) If they are continuous with respective p.d.f.'s $f_i(x_i)$, then

$$E(u(X_1, \dots, X_n)) = \int \cdots \int_{\mathrm{supp}(X_1, \dots, X_n)} u(x_1, \dots, x_n) f_{X_1}(x_1) \cdots f_{X_n}(x_n) \, dx_1 \cdots dx_n.$$

<u>Ex</u> Let X_1 and X_2 be independent random variables with densities $f_{X_1}(x_1) = 3x_1^2$ for $x_1 \in [0,1]$

$$f_{X_2}(x_2) = 5x_2^4 \text{ for } x_2 \in [0,1].$$

independent $\Rightarrow$ $f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$ on $[0,1]^2$.

(a) Find the following.

$$P(.2 < X_1 < .4 \text{ and } .5 \le X_2 < 1) =$$

$$= \int_{.2}^{.4} \int_{.5}^{1} f_{X_1, X_2}(x_1, x_2)\, dx_2\, dx_1$$

$$= \int_{.2}^{.4} \int_{.5}^{1} (3x_1^2)(5x_2^4)\, dx_2\, dx_1$$

$$= \int_{.2}^{.4} 3x_1^2 \cdot x_2^5 \Big|_{x_2=.5}^{x_2=1}\, dx_1$$

$$= \int_{.2}^{.4} 3x_1^2 (1^5 - .5^5)\, dx_1$$

$$= x_1^3 (1^5 - .5^5) \Big|_{x_1=.2}^{x_1=.4} = (.4^3 - .2^3)(1^5 - .5^5)$$

(b) Find the following.

$$E(X_1^6 X_2^7) = \int_0^1 \int_0^1 (x_1^6 x_2^7)(3x_1^2)(5x_2^4)\, dx_2\, dx_1$$

$$= \int_0^1 \int_0^1 15 x_1^8 x_2^{11}\, dx_2\, dx_1$$

$$= 15 \int_0^1 x_1^8\, dx_1 \int_0^1 x_2^{11}\, dx_2$$

$$= 15 \cdot \frac{1}{9} \cdot \frac{1}{12} = \frac{15}{9 \cdot 12}$$

Next we express the expectation of a product of functions of individual independent random variables as the product of expectations of these.

<u>Thm</u> Let $X_1, X_2, \ldots, X_n$ be <u>independent</u> random variables and $u_1(X_1), u_2(X_2), \ldots, u_n(X_n)$ functions of the individual random variables. Then

$$E\left(u_1(X_1)u_2(X_2)\cdots u_n(X_n)\right) = E\left(u_1(X_1)\right)\cdot E\left(u_2(X_2)\right)\cdots E\left(u_n(X_n)\right).$$

<u>Ex</u> Let X_1, X_2, X_3 be independent random variables with respective distributions $b(5,\frac{1}{3}), b(6,\frac{1}{2}), b(7,\frac{3}{4})$.

Find $E(X_1X_2X_3)$.

$$\text{Thm} \Rightarrow E(X_1X_2X_3) = E(X_1)E(X_2)E(X_3)$$

$$= (n_1p_1)(n_2p_2)(n_3p_3)$$

$$= (5\cdot\tfrac{1}{3})(6\cdot\tfrac{1}{2})(7\cdot\tfrac{3}{4})$$

5.4 The Moment-Generating Function Technique for Finding the P.M.F./P.D.F. of a Linear Combination of Independent Random Variables with Known PMFs/PDFs

Recall The moment generating function of a random variable uniquely determines its distribution. In other words, if you know the moment generating function of a random variable, then you can look at a table of moment generating functions, find it, and then match the p.m.f/p.d.f.

In this section, we learn how the m.g.f. changes when you add two random variables or multiply by a scalar, and then use that in combination with tables to find the p.m.f./p.d.f. of a linear combination of random variables.

First recall what the m.g.f. is in the discrete and continuous cases.

Def Let X be a discrete random variable. Then its moment generating function is

$$M(t) := E\left(e^{tX}\right) \overset{\text{Thm}}{=} f(x_1)e^{x_1 t} + f(x_2)e^{x_2 t} + \cdots + f(x_n)e^{x_n t} + \cdots$$

where $\operatorname{supp}(X) = \{x_1, x_2, \ldots, x_n, \ldots\}$.

Here the sum is assumed to converge for all $t \in (-h, h)$ for some positive real h.

Ex If the p.m.f. of X is

x	5	7	9
$f(x)$	1/10	3/10	6/10

then its m.g.f. is $M(t) = \frac{1}{10}e^{5t} + \frac{3}{10}e^{7t} + \frac{6}{10}e^{9t}$

Ex If the m.g.f. of X is $M(t) = \frac{4}{7}e^{100t} + \frac{1}{7}e^{102t} + \frac{2}{7}e^{103t}$, then the p.m.f. is

x	100	102	103
$f(x)$	4/7	1/7	2/7

Ex If X has m.g.f. $\frac{3e^t}{4-e^t}$, then find the distribution of X.

$$\frac{3e^t}{4-e^t} = \frac{3e^t}{4-e^t}\cdot\frac{1/4}{1/4} = \frac{\frac{3}{4}e^t}{1-\frac{1}{4}e^t} = \frac{pe^t}{1-(1-p)e^t}$$

$\Rightarrow X$ has a geometric distribution with $p=\frac{3}{4}$.

Def Let X be a continuous random variable. Then its moment generating function is

$$M(t) := \int_{-\infty}^{\infty} e^{tx} f(x)\,dx .$$

Here the integral is assumed to converge for all $t\in(-h,h)$ for some positive real h.

Ex If X has m.g.f. $M(t)=\frac{2}{2-t}$, find its density and variance.

$$\frac{2}{2-t} = \frac{2}{2-t}\cdot\frac{1/2}{1/2} = \frac{1}{1-\frac{1}{2}t} = \frac{1}{1-\theta t}$$

$\Rightarrow$ exponential

$f(x) = \frac{1}{1/2}e^{-x/(1/2)}$ $\quad x\in[0,\infty)$

$\mathrm{Var}(X) = \theta^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$

Ex Use the definition of moment generating function for a continuous random variable to find the moment generating function of a random variable X uniformly distributed on $[a,b]$.

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx = \int_a^b e^{tx} \cdot \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a}\left(\frac{1}{t} e^{tx}\Big|_{x=a}^{x=b}\right)$$

$$= \frac{e^{tb} - e^{ta}}{t(b-a)}$$

We next learn how to find the m.g.f. of a linear combination.

Thm Let $X_1, X_2, \ldots, X_n$ be independent random variables with respective moment generating functions $M_{X_i}(t)$. Let $a_1, a_2, \ldots, a_n$ be real constants. Let $Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$.

Then
$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t).$$

Ex Let X_1 be an exponential random variable with mean 3, and let X_2 be $N(2,9)$. Find the m.g.f. of $7X_1 + 10X_2 =: Y$.

From the tables,

$$M_{X_1}(t) = \frac{1}{1-\theta t} = \frac{1}{1-3t},$$

$$M_{X_2}(t) = \exp(\mu t + \sigma^2 t^2/2) = \exp(2t + 9t^2/2)$$

$$\text{Thm} \Rightarrow M_Y(t) = M_{X_1}(a_1 t) \cdot M_{X_2}(a_2 t)$$

$$= M_{X_1}(7t) \cdot M_{X_2}(10t)$$

$$= \frac{1}{1-3\cdot 7t} \cdot \exp(2\cdot 10t + 9\cdot 100t^2/2)$$

$$= \frac{1}{1-21t} \cdot \exp(20t + 900t^2/2).$$

Next we do a discrete example that does not refer to the tables.

Ex Consider a fair coin with sides labelled 1 and 2 and consider a fair 3-sided die. Suppose we simultaneously flip the coin and roll the die. Let X_1 be the number on the coin and X_2 the number on the die, and $Y := X_1 + X_2$.

Find the p.m.f. of Y using the moment generating function technique, i.e. find the m.g.f. of Y using the Theorem, and then determine the p.m.f. from the m.g.f..

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

$$= \left(\frac{1}{2}e^{1t} + \frac{1}{2}e^{2t}\right) \cdot \left(\frac{1}{3}e^{1t} + \frac{1}{3}e^{2t} + \frac{1}{3}e^{3t}\right)$$

$$= \frac{1}{6}\left(e^{t}\left(e^{t} + e^{2t} + e^{3t}\right) + e^{2t}\left(e^{t} + e^{2t} + e^{3t}\right)\right)$$

$$= \frac{1}{6}\left(e^{2t} + e^{3t} + e^{4t} + e^{3t} + e^{4t} + e^{5t}\right)$$

$$= \frac{1}{6}\left(e^{2t} + 2e^{3t} + 2e^{4t} + e^{5t}\right)$$

$$= \frac{1}{6}e^{2t} + \frac{1}{3}e^{3t} + \frac{1}{3}e^{4t} + \frac{1}{6}e^{5t}$$

$\Rightarrow$ The p.m.f. is

y	2	3	4	5
$f_Y(y)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

Ex Suppose 3 people are standing at the entrance to Millenium Park in Chicago on Michigan Avenue. Each wants to hail a cab to go to another part of the city. Suppose empty cabs drive by according to an exponential distribution with mean 4 minutes.
What is the probability that all 3 people have departed within 10 minutes? Assume independence.

$$\text{wait time} = Y = \underbrace{X_1 + X_2 + X_3}_{\text{exp. w/ mean 4.}}$$

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot M_{X_3}(t)$$

$$= \frac{1}{1-\theta t} \cdot \frac{1}{1-\theta t} \cdot \frac{1}{1-\theta t}$$

$$= \frac{1}{1-4t} \cdot \frac{1}{1-4t} \cdot \frac{1}{1-4t}$$

$$= \frac{1}{(1-4t)^3} = \frac{1}{(1-\theta t)^\alpha}$$

$\Rightarrow Y$ has a gamma distribution with $\alpha = \text{shape} = 3$
$\theta = \text{scale} = 4.$

$$P(Y \leq 10) = \int_0^{10} \frac{1}{\Gamma(3)4^3} x^{3-1} e^{-x/4} dx = \text{pgamma}(10, \text{shape}=3, \text{scale}=4)$$

$$= .1912$$

Thm Let $X_1, X_2, \ldots, X_n$ be independent chi-square random variables with $r_1, r_2, \ldots, r_n$ respective degrees of freedom. Then the sum $Y = X_1 + X_2 + \cdots + X_n$ is chi-square with $r_1 + r_2 + \cdots + r_n$ degrees of freedom.

Ex Let X_1, X_2, X_3 be chi-square with 7, 8, 9 degrees of freedom. What is the density of $X_1 + X_2 + X_3$?

$r = 7 + 8 + 9 = 24.$

$$\text{density} = \frac{1}{\Gamma(r/2)\, 2^{r/2}} x^{r/2 - 1} \cdot e^{-x/2} = \frac{1}{11!\, 2^{12}} x^{11} \cdot e^{-x/2}$$

for $x \in [0, \infty)$.

Thm (i) Let Z be a standard normal random variable. Then Z^2 is $\chi^2(1)$.

(ii) If $Z_1, Z_2, \ldots, Z_n$ are all standard normals, then $Z_1^2 + Z_2^2 + \cdots + Z_n^2$ is $\chi^2(n)$.

5.5 Distributions of Functions of Normal Random Variables

Recall Random variables $X_1, X_2, \ldots, X_n$ are said to be independent and identically distributed or i.i.d. if

1) their joint p.m.f./p.d.f. is the product of the individual p.m.f.'s/p.d.f.'s

and

2) the individual p.m.f.'s/p.d.f.'s are all the same (just in $x_1, x_2, \ldots, x_n$).

If $X_1, X_2, \ldots, X_n$ are i.i.d., then we synonymously say they are a random sample of size n from the common distribution.

Ex Let X_1, X_2, X_3 be a random sample from an exponential distribution with mean 2. Then the joint density is $f(x_1, x_2, x_3) = \left(\frac{1}{2}e^{-\frac{1}{2}x_1}\right)\left(\frac{1}{2}e^{-\frac{1}{2}x_2}\right)\left(\frac{1}{2}e^{-\frac{1}{2}x_3}\right)$ for $(x_1, x_2, x_3) \in [0, \infty)^3$.

<u>Def</u> Any function $u(X_1, X_2, \dots, X_n)$ of a random sample $X_1, X_2, \dots, X_n$ from a distribution is called a <u>statistic</u>.

<u>Ex</u> Let $X_1, X_2, \dots, X_n$ be a random sample from a distribution. Then the <u>sample mean</u>

$$\overline{X} := \frac{1}{n} \sum_{i=1}^{n} X_i$$

is an example of a statistic. The <u>sample variance</u>

$$S^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

is also an example of a statistic.

<u>Rem</u> If a population is known (or assumed) to be normal, then one would like to estimate the parameters μ and σ^2, or test conjectures about them.
The sample mean $\overline{X}$ and sample variance S^2 from above can be used for this purpose. Today we learn some theorems about statistics from normal populations.

First we begin with a theorem that does <u>not</u> assume all $X_1, X_2, \ldots, X_n$ have the same normal distribution.

<u>Thm</u> Let $X_1, X_2, \ldots, X_n$ be independent normal random variables with respective means $\mu_1, \mu_2, \ldots, \mu_n$ and respective variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$.

Then: 1) The sum $X_1 + X_2 + \cdots + X_n$ is also <u>normal</u> with mean $\mu_1 + \mu_2 + \cdots + \mu_n$ and variance $\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$.

2) If $c_1 \in \mathbb{R}$, then $c_1 X_1$ is also <u>normal</u> with mean $c_1 \mu_1$ and variance $c_1^2 \sigma_1^2$.

3) If $c_1, c_2, \ldots, c_n \in \mathbb{R}$, then the linear combination

$$c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$$

is also <u>normal</u>, with mean

$$c_1 \mu_1 + c_2 \mu_2 + \cdots + c_n \mu_n$$

and variance $c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 + \cdots + c_n^2 \sigma_n^2$.

Ex If X_1 is $N(5,9)$ and X_2 is $N(7,36)$,
then $4X_1 - 13X_2$ is $N(4\cdot 5 - 13\cdot 7,\ 4^2\cdot 9 + (-13)^2\cdot 36)$

is $N(20-91,\ 144+6084)$

is $N(-71,\ 6228)$.

The standard deviation of $4X_1 - 13X_2$ is $\sqrt{6228}$.

Ex A certain electronic system works if either (or both) of two resistors are functional. The life span of resistor 1 is the random variable X_1 (measured in months), and is $N(70,25)$. The life span of resistor 2 is the random variable X_2 (measured in months) and is $N(80,144)$.

(a) What is the probability that resistor 2 lasts more than 7 months longer than resistor 1?

$$P(X_2 > X_1 + 7) = P(X_2 - X_1 > 7)$$

X_1 and X_2 are independent

$$= P\left(\frac{(X_2 - X_1) - (80-70)}{\sqrt{144+(-1)^2\cdot 25}} > \frac{7-(80-70)}{\sqrt{144+(-1)^2\cdot 25}}\right)$$

$$\approx P(Z > -.2308)$$

$= \text{pnorm}(-.2308, \text{ lower.tail} = \text{FALSE})$

$= .59126\ldots$

(b) Find the probability that the electronic system still works after 90 months.

$$P(X_1 > 90 \text{ or } X_2 > 90) = P((X_1 > 90) \cup (X_2 > 90))$$

$$= P(X_1 > 90) + P(X_2 > 90) - P((X_1 > 90) \cap (X_2 > 90))$$

$$= P(X_1 > 90) + P(X_2 > 90) - \underbrace{P(X_1 > 90)\,P(X_2 > 90)}_{\text{this is because } X_1 \text{ and } X_2 \text{ are independent.}}$$

$= \text{pnorm}(90, \text{ mean} = 70, \text{ sd} = 5, \text{ lower.tail} = \text{FALSE})$
$+ \text{pnorm}(90, \text{ mean} = 80, \text{ sd} = 12, \text{ lower.tail} = \text{FALSE})$
$-$ (product of previous two terms)

$= .20235\ldots$

If all the random variables $X_1, X_2, \ldots, X_n$ have the same normal distribution in the previous Thm, then we get the following result about the sample mean.

Cor Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2. Then the sample mean $\overline{X} := \frac{1}{n} \sum_{i=1}^{n} X_i$ is normal with mean μ and variance $\frac{\sigma^2}{n}$.

Rem This means for large n: the sample mean has a high probability of being close to μ, since the dispersion is small (σ^2 gets divided by n).

Ex Let $X_1, X_2, \ldots, X_{100}$ be a random sample from a normal distribution $N(50, 49)$. Then

$$P(49 < \overline{X} < 51) = \text{pnorm}(51, \text{mean}=50, \text{sd}=\tfrac{7}{10}) - \text{pnorm}(49, \text{mean}=50, \text{sd}=\tfrac{7}{10})$$

$$\approx .8469$$

Very high probability to be in small interval $(49, 51)$!

In statistical applications we will use the following.

<u>Thm</u> Let $X_1, X_2, \dots, X_n$ be a random sample from a normal distribution $N(\mu, \sigma^2)$.

Then: 1) the sample mean $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$

and sample variance $S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

are independent, and

2) the following multiple of the sample variance

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2}$$

is $\chi^2(n-1)$.

<u>Ex</u> If $X_1, X_2, \dots, X_{15}$ are a random sample from a normal distribution $N(7, 25)$, find

$$P\left(625 \le \sum_{i=1}^{15}(X_i - 7)^2 \le 875\right) \text{ and } P\left(625 \le \sum_{i=1}^{15}(X_i - \overline{X})^2 \le 875\right).$$

Notice difference! The first is the mean of the population, the second is the sample mean!

First, observe that $\dfrac{\sum_{i=1}^{15}(X_i-7)^2}{25} = \sum_{i=1}^{15}\left(\dfrac{X_i-7}{5}\right)^2$

is $\chi^2(15)$ because it is the sum of 15 squares of standard normals (see Section 5.4).

Second, observe that $\dfrac{\sum_{i=1}^{15}(X_i-\bar{X})^2}{25} = \dfrac{(n-1)S^2}{\sigma^2}$

is $\chi^2(n-1)$, i.e. is $\chi^2(14)$.

Thus,

$$P\left(625 \le \sum_{i=1}^{15}(X_i-7)^2 \le 875\right) = P\left(\frac{625}{25} \le \sum_{i=1}^{15}\left(\frac{X_i-7}{5}\right)^2 \le \frac{875}{25}\right)$$

$$= \text{pchisq}\left(\frac{875}{25}, df=15\right) - \text{pchisq}\left(\frac{625}{25}, df=15\right)$$

$$\approx .0475$$

and

$$P\left(625 \le \sum_{i=1}^{15}(X_i-\bar{X})^2 \le 875\right) = P\left(\frac{625}{25} \le \sum_{i=1}^{15}\frac{(X_i-\bar{X})^2}{25} \le \frac{875}{25}\right)$$

$$= \text{pchisq}\left(\frac{875}{25}, df=14\right) - \text{pchisq}\left(\frac{625}{25}, df=14\right)$$

$$\approx .0331$$

Some of the most important inferences in statistics use Student's t distribution.

<u>Thm</u> Suppose Z is a standard normal random variable and U is $\chi^2(r)$. Suppose also that Z and U are independent. Then the random variable

$$T := \frac{Z}{\sqrt{U/r}}$$

has pdf

$$f(t) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r}\,\Gamma\left(\frac{r}{2}\right)} \cdot \frac{1}{\left(1+\frac{t^2}{r}\right)^{\frac{r+1}{2}}} \qquad t \in \mathbb{R}.$$

This density is called a <u>t-distribution with r degrees of freedom</u>. We say the distribution is <u>$t(r)$</u>.

<u>Rem</u> The t-distributions look like the standard normal, but have thicker tails than the standard normal, and consequently are slightly lower than $\frac{1}{\sqrt{2\pi}} \approx .399$ at 0.

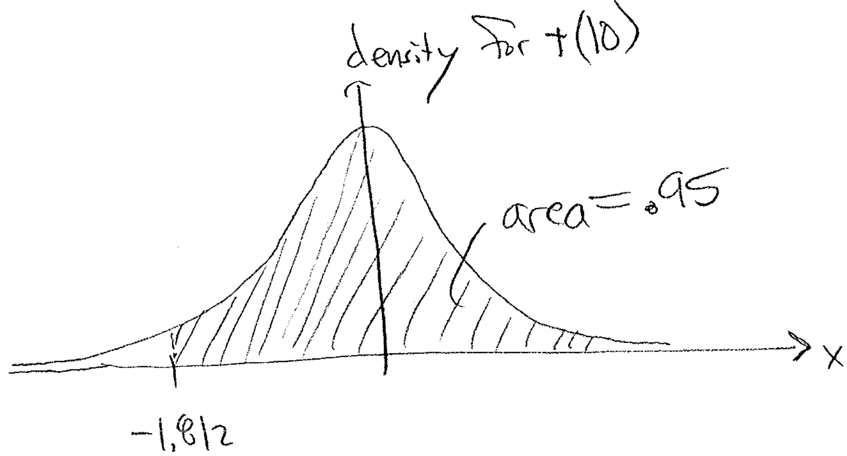
Ex Type into R:

```
curve(dnorm(x), -4,4, col="red")
curve( dt(x, df=1), -4,4, add=TRUE)
curve(dt(x, df=2), -4,4, add=TRUE)
curve(dt(x, df=3), -4,4, add=TRUE)
curve(dt(x, df=10), -4,4, add=TRUE)
```

Notation $t_\alpha(r)$ is the number such that the area under the $t(r)$ density to the right of $t_\alpha(r)$ is α.

Ex Find $t_{.95}(10)$ and draw a picture.

qt(.95, df=10, lower.tail=FALSE) ≈ −1.812



Prop Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution $N(\mu, \sigma^2)$. Then $\dfrac{\bar{X} - \mu}{S/\sqrt{n}}$ has a Student's t-distribution with $r = n-1$ degrees of freedom.

Rem We will use this t-distribution in Section 6.2 to construct confidence intervals for an unknown mean μ of a normal distribution.

Ex Let $X_1, X_2, \ldots, X_5$ be a random sample from a normal distribution $N(10, 64)$. If

$$T = \frac{\bar{X} - 10}{\left(\frac{1}{4}\sum_{i=1}^{5}(X_i - \bar{X})^2\right)^{\frac{1}{2}} / \sqrt{5}},$$

Find $P(T \leq 1)$.

This is $pt(1, df=4) \approx .813$

5.6 The Central Limit Theorem

In the previous section we saw that when $X_1, X_2, \dots, X_n$ is a random sample from a normal distribution $N(\mu, \sigma^2)$, then $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is normal with same mean μ but smaler variance σ^2/n. The main point of this section is to see that for a random sample from any distribution (continuous or discrete), for large enough n, the sample mean $\overline{X}$ is approximately normal with mean μ and variance σ^2/n.

But first we need a small lemma, which implies translations of normal random variables are also normal.

Lemma Suppose X is normal with mean μ and variance σ^2. Let $a, b \in \mathbb{R}$. Then the random variable $aX + b$ is also normal with mean $a\mu + b$ and variance $a^2\sigma^2$.

Ex If X is $N(2,25)$, find the distribution of $Y=-4X-7$

By the lemma, Y is normal with mean $-4\cdot 2-7=-15$ and variance $(-4)^2\cdot 25=400$. $\Rightarrow Y \sim N(-15,400)$.

Rem Last time we saw that when $X_1, X_2, \dots, X_n$ is a random sample from a normal distribution $N(\mu,\sigma^2)$, then the sample mean $\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ is normal with mean μ and variance σ^2/n. So its standardization

$$\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}$$

is standard normal, i.e. $N(0,1)$. We next see in the Central Limit Theorem, without the assumption of normality on the random sample, the "standardization" of the sample mean tends to the standard normal, as the size of the random sample increases.

Thm (Central Limit Theorem) Let $X_1, X_2, \ldots, X_n, X_{n+1}, \ldots$ be a sequence of i.i.d. random variables, with common mean μ and common variance σ^2. Let $\overline{X}_n$ be the sample mean of the first n random variables, i.e. $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$.

Then the distribution of the "standardization" of $\overline{X}_n$

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$$

tends to the standard normal as $n \to \infty$. That is,

$$\lim_{n\to\infty} P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le w\right) = \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz.$$

The convergence is uniform in w.

Cor Let $X_1, X_2, \ldots, X_n$ be a random sample from a common distribution with mean μ and variance σ^2. Suppose n is large. Then

1) The sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is approximately $N(\mu, \sigma^2/n)$.

2) The sum $\sum_{i=1}^{n} X_i$ is approximately $N(n\mu, n\sigma^2)$.

Pr 1) By CLT, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ is approximately $N(0,1)$.

Lemma $\Rightarrow$ $\underbrace{\left(\frac{\sigma}{\sqrt{n}}\right)\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)+\mu}_{= \bar{X}}$ is approximately $N\left(\frac{\sigma}{\sqrt{n}}\cdot 0+\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) = N\left(\mu, \frac{\sigma^2}{n}\right)$.

2) By 1), $\frac{1}{n}\sum_{i=1}^{n} X_i$ is approximately $N\left(\mu, \frac{\sigma^2}{n}\right)$,

so Lemma $\Rightarrow$ $\underbrace{n\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)}_{= \sum_{i=1}^{n} X_i}$ is approximately $N\left(n\mu, n^2\frac{\sigma^2}{n}\right) = N(n\mu, n\sigma^2)$

Ex Let $\overline{X}$ be the mean of a random sample of size 15 from the uniform distribution $U(4,10)$. Approximate $P(6.5 \le \overline{X} \le 7.5)$ using the CLT and using part 1) of the corollary.

The mean of the common distribution is

$$\mu = \frac{a+b}{2} = \frac{4+10}{2} = \frac{14}{2} = 7$$

and the variance of the common distribution is

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(10-4)^2}{12} = \frac{36}{12} = 3.$$

Using CLT: $P(6.5 \le \overline{X} \le 7.5) = P\left(\frac{6.5-\mu}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{7.5-\mu}{\sigma/\sqrt{n}}\right)$

$$\overset{CLT}{\approx} P\left(\frac{6.5-7}{\sqrt{3}/\sqrt{15}} \le Z \le \frac{7.5-7}{\sqrt{3}/\sqrt{15}}\right)$$

$$= \text{pnorm}(.5/(\sqrt{3}/\sqrt{15})) - \text{pnorm}(-.5/(\sqrt{3}/\sqrt{15}))$$

Using Cor: $\overline{X} \sim N(\mu, \sigma^2/n) = N(7, 3/15)$, so

$$P(6.5 \le \overline{X} \le 7.5) = \text{pnorm}(7.5, \text{mean}=7, \text{sd}=\sqrt{3/15})$$
$$- \text{pnorm}(6.5, \text{mean}=7, \text{sd}=\sqrt{3/15})$$

In both methods, we get approximation .73644, which is a high probability for such a small interval! Using sample size 40, we get a new random variable $\overline{X}_{40}$ which has $P(6.5 \leq \overline{X}_{40} \leq 7.5) \approx .93211$

Ex Let $X_1, X_2, \ldots, X_{30}$ be a random sample from a Poisson distribution with mean 2.

Let Y be the sum $X_1 + X_2 + \cdots + X_{30}$.

Approximate $P(Y \leq 70)$ using the corollary and compare to the exact value.

Cor $\Rightarrow$ Y is approximately $N(n\mu, n\sigma^2) = N(30 \cdot 2, 30 \cdot 2)$
$= N(60, 60)$.

$$\Rightarrow P(Y \leq 70) \overset{\text{Cor}}{\approx} \text{pnorm}(70, \text{mean}=60, \text{sd}=\text{sqrt}(60))$$
$$\approx .90164\ldots$$

For the exact value, recall from 5.4 #4 that Y is Poisson with mean $\underbrace{2+2+\cdots+2}_{30 \text{ times}} = 60$.

$$\Rightarrow P(Y \leq 70) = \text{ppois}(70, \text{lambda}=60)$$
$$= .9098\ldots \quad \text{close to approximation!}$$