

# Assignment 1 - Thomas Boyko / Timothy Markovic

1. Let  $X$  be a random variable with a density function given by

$$f_X(x) = \begin{cases} \frac{3}{2}x^2, & -1 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the density function of  $Y = 3 - X$ .

First we find our support for  $Y$ . From  $Y = 3 - X$  and  $X \in [-1, 1]$ , we get  $Y \in [2, 4]$ . And we can quickly find that  $x = 3 - y$ .

Next we find  $F_Y(y)$ .

$$\begin{aligned} F_Y(y) &= F_X(x) \left| \frac{dx}{dy} \right| \\ &= F_X(3 - y) I(-1 \leq x \leq 1) \left| \frac{d}{dy}(3 - y) \right| \\ &= \frac{3}{2}(3 - y)^2 I(-1 \leq x \leq 1) |-1| \\ &= \frac{3}{2}(3 - y)^2 I(-1 \leq x \leq 1) \end{aligned}$$

This gives our PDF:

$$f_Y(y) = I(2 \leq y \leq 4)(3 - y)^2$$

(b) Find the density function of  $Y = X^2$ .

Again we start with our support. We see  $Y$  is maximised at  $x = 1$  and minimised at  $x = 0$ . So our support for  $Y$  is  $[0, 1]$ .

We must split into two cases since  $Y$  is two to one.

$$\begin{aligned} f_Y(y) &= f_X(-\sqrt{y}) \left| \frac{d}{dy}(-\sqrt{y}) \right| + f_X(\sqrt{y}) \left| \frac{d}{dy}\sqrt{y} \right| \\ &= \frac{3}{2}y\left(\frac{1}{2\sqrt{y}}\right) + \frac{3}{2}y\left(\frac{1}{2\sqrt{y}}\right) \\ &= \frac{3}{2}\sqrt{y} \end{aligned}$$

We find the PDF (with support):

$$f_Y(y) = \frac{3}{4}\sqrt{y}I(0 \leq y \leq 1).$$

2. Assume that  $X$  has a beta distribution with parameters  $\alpha$  and  $\beta$ .

(a) Find the density function of  $Y = 1 - X$ .

We again start with our support.  $x \in [0, 1]$  since  $X \sim \text{Beta}(\alpha, \beta)$ , so  $y \in [0, 1]$ .

Next, we find  $x$  as a function of  $y$ ;  $x = 1 - y$ .

Finally, we write our pdf for  $X$ .

$$f_X(x) = B(\alpha, \beta) I(0 \leq x \leq 1) x^{\alpha-1} (1 - x)^{\beta-1}.$$

Where  $B(\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ .

With this we can find  $f_Y(y)$ .

$$\begin{aligned} f_Y(y) &= f_X(x) \left| \frac{dx}{dy} \right| \\ &= f_X(1 - y) \left| \frac{d}{dy}(1 - y) \right| \\ &= B(\alpha, \beta) I(0 \leq 1 - y \leq 1) (1 - y)^{\alpha-1} (1 - (1 - y))^{\beta-1} |-1| \\ &= B(\alpha, \beta) I(0 \leq y \leq 1) (1 - y)^{\alpha-1} y^{\beta-1} \end{aligned}$$

- (b) Identify the density of  $Y$  as a well-known distribution. Be sure to identify any parameter values.

$Y$  takes the density function of a beta distribution with  $\alpha, \beta$  opposite from  $X$ .

3. The Weibull density function is given by:

$$f_X(x) = \begin{cases} \frac{1}{\alpha} m x^{m-1} e^{-\frac{x^m}{\alpha}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

where  $\alpha$  and  $m$  are positive constants. This density function is often used as a model for the lengths of life of physical systems. Suppose  $X$  has the Weibull density just given. Find:

- (a) The density function of  $Y = X^m$ .

Begin by finding  $x$  as a function of  $y$ . Since  $x$  is positive, we can take root  $m$  of  $x^m$  and get  $x = y^{\frac{1}{m}}$ .

Now we can find  $f_Y(y)$ .

$$\begin{aligned} f_Y(y) &= f_X(x) \left| \frac{dx}{dy} \right| \\ &= f_X(y^{\frac{1}{m}}) \left| \frac{d}{dy} y^{\frac{1}{m}} \right| \\ &= \frac{1}{\alpha} m (y^{\frac{1}{m}})^{m-1} e^{-\frac{y}{\alpha}} \left| \frac{y^{\frac{1}{m}-1}}{m} \right| \\ &= \frac{1}{\alpha} (y^{\frac{m-1}{m}}) e^{-\frac{y}{\alpha}} y^{\frac{1-m}{m}} \\ &= \frac{1}{\alpha} (y^{\frac{m-1+1-m}{m}}) e^{-\frac{y}{\alpha}} \\ &= \frac{1}{\alpha} (y^{\frac{0}{m}}) e^{-\frac{y}{\alpha}} \\ &= \frac{1}{\alpha} e^{-\frac{y}{\alpha}}. \end{aligned}$$

Which after simplification gives us an exponential distribution with mean  $\alpha$ .

- (b)  $E(X^k)$  for any positive integer  $k$ .

$$\begin{aligned} E(X^k) &= \int_0^\infty x^k \frac{m}{\alpha} x^{m-1} \exp\left(-\frac{x^m}{\alpha}\right) dx \\ &= \frac{m}{\alpha} \int_0^\infty x^{k+m-1} \exp\left(-\frac{x^m}{\alpha}\right) dx \\ &= \frac{m}{\alpha} \int_0^\infty (t\alpha)^{\frac{k+m-1}{m}} e^{-t} (t\alpha)^{\frac{1}{m}-1} dt & t = x^m/\alpha \\ &= \alpha^{1-1+\frac{k}{m}+\frac{1}{m}-\frac{1}{m}} \int_0^\infty t^{1-1+\frac{1}{m}-\frac{1}{m}+\frac{k}{m}} e^{-t} dt & x = (t\alpha)^{\frac{1}{m}} \\ &= \alpha^{\frac{k}{m}} \int_0^\infty t^{\frac{k}{m}+1-1} e^{-t} dt \\ &= \alpha^{\frac{k}{m}} \Gamma\left(\frac{k}{m} + 1\right) \end{aligned}$$

4.  $X$  is a uniformly distributed random variable between 0 and 1.

(a) Find the probability density function of  $Y = -\lambda \ln(X)$ ,  $\lambda > 0$ .

We begin with our support for  $Y$ . As  $X$  approaches 0,  $Y$  will approach infinity. And when  $x = 1$ ,  $y = 0$ . So our support for  $Y$  is  $Y \in (0, \infty)$ .

Recall the uniform distribution between 0 and 1 is just:  $f_X(x) = I(0 \leq x \leq 1)$ .  $y = -\lambda \ln x$  so:

$$\begin{aligned} e^y &= e^{-\lambda \ln x} \\ &= e^{\ln(x^{-\lambda})} \\ &= x^{-\lambda} \\ x &= e^{-\frac{y}{\lambda}}. \end{aligned}$$

Now that we have  $x$  as a function of  $y$ , we can find the PDF:

$$\begin{aligned} f_Y(y) &= f_X(x) \left| \frac{dx}{dy} \right| I(0 \leq x \leq 1) \\ &= f_X(e^{-\frac{y}{\lambda}}) \left| \frac{d}{dy} e^{-\frac{y}{\lambda}} \right| I(y > 0) \\ &= 1 \left| -\frac{1}{\lambda} e^{-\frac{y}{\lambda}} \right| I(y > 0) \\ &= \frac{1}{\lambda} e^{-\frac{y}{\lambda}} I(y > 0) \end{aligned}$$

(b) Find the expected value and standard deviation of  $Y$ . ( $\mu_Y = E[Y]$ ,  $\sigma_Y = SD[Y]$ ).

Since  $Y$  is an exponential distribution, we can say that its mean and standard deviation are both equal to  $\lambda$ .

5. The lifetime of an electronic component in an HDTV is a random variable that can be modeled by the exponential distribution with a mean lifetime  $\beta$ . Two components,  $X_1$  and  $X_2$ , are randomly chosen and operated until failure. At that point, the lifetime of each component is observed. The mean lifetime of these two components is  $\bar{X} = \frac{X_1 + X_2}{2}$ .

(a) Find the probability density function of  $X$ . Use the MGF method.

We begin with the exponential distribution's MGF, giving us  $M_{X_1}(t) = M_{X_2}(t) = (1 - \beta t)^{-1}$ . Also note that  $X_1$  and  $X_2$  are independent since they are random samples.

Then:

$$\begin{aligned} M_{\bar{X}}(t) &= E[e^{\bar{X}t}] \\ &= E\left[\exp\left(\frac{X_1 + X_2}{2}t\right)\right] \\ &= E\left[\exp\left(\frac{X_1 t}{2}\right) \exp\left(\frac{X_2 t}{2}\right)\right] \\ &= E\left[\exp\left(X_1 \frac{t}{2}\right)\right] E\left[\exp\left(X_2 \frac{t}{2}\right)\right] && \text{By independence of } X_1, X_2. \\ &= M_{X_1}\left(\frac{t}{2}\right) M_{X_2}\left(\frac{t}{2}\right) \\ &= (1 - \beta \frac{t}{2})^{-1} (1 - \beta \frac{t}{2})^{-1} \\ &= (1 - \beta \frac{t}{2})^{-2}. \end{aligned}$$

We recognise the MGF of a gamma distribution,  $\bar{X} \sim \text{Gamma}(\beta = \frac{\beta}{2}, \alpha = 2)$ . From this we can get our pdf:

$$f_{\bar{X}}(\bar{x}) = \frac{4\bar{x}}{\beta^2} \exp\left(-\frac{2\bar{x}}{\beta}\right).$$

- (b) If the mean lifetime of the electronic component is two years ( $\beta = 2$ ), what is the probability that the mean lifetime of the two components tested will be more than 3.0 years? Show all relevant work.

We substitute  $\beta = 2$  into our above distribution, giving us  $Gamma(1, 2)$ . We want  $P(X > 3) = 1 - P(X \leq 3)$ . We can use  $R$ :

```
> 1-pgamma(3,2,1)
[1] 0.1991483
```

6. Let  $X_1, X_2, \dots, X_{10}$  represent a sample of size 10 taken from a normal distribution with  $\mu = 0$  and  $\sigma^2 = 1$ . Defining  $U$  as

$$U = \sum_{n=1}^{10} X_i^2.$$

- (a) Find the distribution of  $U$ . In addition, provide the mean and standard deviation of  $U$ :  $(\mu_U, \sigma_U)$

By Theorem 7.2, we can say that  $U \sim \chi_{df=10}^2$ .

And from our distribution chart,  $\mu_U = \nu = 10$ , and  $\sigma_U = \sqrt{2\nu} = \sqrt{20}$ .

- (b) From your answer in (a), find  $P(12 \leq U \leq 15)$ .

We simply calculate in  $R$ :

```
> pchisq(15,10)-pchisq(12,10)
[1] 0.1529946
```

- (c) Find  $\mu_U$  the median of  $U$ . Note, the median is a value such that  $P(U \leq \mu_U) = 0.50$ .

We find that in  $R$ :

```
> qchisq(0.5,10)
[1] 9.341818
```

7. Let  $X$  and  $Y$  be two independent random variables, where  $X \sim Normal(0, 1)$  and  $Y \sim Normal(0, 1)$ . A random variable  $Z$  is defined as  $Z = \frac{X}{Y}$ .

- (a) Using any of the methods discussed in class, find the probability distribution function of  $Z$ , or  $f_Z(z)$ .

We begin by finding  $X, Y$  as inverse functions. We have  $X = ZY$  and  $Y = \frac{X}{Z}$ . Since  $X, Y \in \mathbb{R}$ ,  $Z \in \mathbb{R}$ .

Recall our distributions for  $X, Y$ :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

And since  $X \perp Y$ :

$$f_{XY}(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right)$$

Now we find the pdf of  $Z$ .

$$\begin{aligned}
f_{XZ}(x, z) &= f_{XY}(x, y) \left| \frac{dx}{dz} \right| \\
&= f_{XY}(yz, y) \left| \frac{d}{dz} zy \right| \\
&= \frac{|y|}{2\pi} \exp\left(-\frac{y^2 z^2 + y^2}{2}\right) \\
&= \frac{|y|}{2\pi} \exp\left(-\frac{y^2(z^2 + 1)}{2}\right) \\
&= \frac{|y|}{2\pi} \exp\left(-\frac{y^2(z^2 + 1)}{2}\right).
\end{aligned}$$

Now we must integrate  $Y$  out of the pdf over the support of  $y$ ,  $(-\infty, \infty)$ :

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} \frac{|y|}{2\pi} \exp\left(-\frac{y^2(z^2 + 1)}{2}\right) dy \\
&= \int_{-\infty}^0 -\frac{y}{2\pi} \exp\left(-y^2 \frac{z^2 + 1}{2}\right) dy + \int_0^{\infty} \frac{y}{2\pi} \exp\left(-y^2 \frac{z^2 + 1}{2}\right) dy \\
&= \frac{1}{2\pi} \left( \int_{y=-\infty}^{y=0} \frac{-1}{2} \exp\left(-u \frac{z^2 + 1}{2}\right) dy + \int_{y=0}^{y=\infty} \frac{1}{2} \exp\left(-u \frac{z^2 + 1}{2}\right) dy \right) \\
&= \frac{1}{4\pi} \left( \int_0^{\infty} \exp\left(-u \frac{z^2 + 1}{2}\right) dy + \int_0^{\infty} \exp\left(-u \frac{z^2 + 1}{2}\right) dx \right) \\
&= \frac{1}{2\pi} \left( -\frac{2}{z^2 + 1} \right) \exp\left(-u \frac{z^2 + 1}{2}\right) \Big|_{u=0}^{\infty} \\
&= -\frac{1}{\pi(z^2 + 1)}(0 - 1) \\
&= \frac{1}{\pi(z^2 + 1)}
\end{aligned}$$

With support:  $z \in \mathbb{R}$ .

(Used the substitution  $y^2 = u$ )

(b) Find  $\mu_z = E[Z]$ .

We use the formula  $\mu_Z = \int_{\text{all } Z} z f_Z(z) dz$ .

$$\begin{aligned}
\mu_Z &= \int_{-\infty}^{\infty} \frac{z}{\pi(z^2 + 1)} dz & u = z^2 + 1, dz = \frac{du}{2z} \\
&= 0 & \text{Since } f_Z(z) \text{ is odd.}
\end{aligned}$$

(c) What do you notice about the  $\sigma_Z = \text{Var}(Z)$ ? Anything peculiar? Comment.

The variance of  $Z$  is given by:

$$\begin{aligned}
\sigma_Z &= E[Z^2] - \mu_Z \\
&= \int_{-\infty}^{\infty} \frac{z^2}{2\pi(z^2 + 1)} dz & z = \tan \theta \\
&= \frac{1}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \tan^2 \theta d\theta \\
&= \frac{1}{2\pi} (\tan \theta - \theta) \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}
\end{aligned}$$

$\tan z$  is undefined at both of our endpoints, so the variance does not exist.

8. Let  $X_1, X_2, \dots, X_n$  represent a random sample of observations taken from an exponentially distributed population with scale parameter  $\beta$ . Let  $Y = X_1 + X_2 + \dots + X_n$ . Assuming the  $X_i$ 's are independent random variables, identify the distribution of the random variable  $Z = \frac{2Y}{\beta}$ .

We begin by finding the distribution of  $Y$ . Start by noting that:  $M_{X_i}(t) = (1 - \beta t)^{-1}$ .

Then

$$\begin{aligned}
 M_Y(t) &= E[e^{tY}] \\
 &= E\left[\exp\left(t \sum_{i=1}^n X_i\right)\right] \\
 &= E\left[\prod_{i=1}^n \exp(tX_i)\right] \\
 &= \prod_{i=1}^n E[\exp(tX_i)] && \text{By Independence of all } X_i \\
 &= \prod_{i=1}^n M_{X_i}(t) \\
 &= M_{X_i}(t)^n \\
 &= (1 - \beta t)^{-n}
 \end{aligned}$$

This is the MGF of a gamma distribution, so  $Y \sim \text{Gamma}(\beta = \beta, \alpha = n)$ .

Now we use the method of transformations to find the distribution of  $Z$ . Since  $Z = \frac{2Y}{\beta}$ ,  $Y = \frac{\beta Z}{2}$ . And from  $Y > 0$  since  $Y$  is a gamma distribution, we can say that  $Z > 0$ .

Now we are ready to find the pdf of  $Z$ .

We start with the distribution of  $Y$ :

$$f_Y(y) = \frac{1}{\Gamma(n)\beta^n} y^{n-1} e^{-\frac{y}{\beta}}.$$

$$\begin{aligned}
 f_Z(z) &= f_Y(y) \left| \frac{dy}{dz} \right| \\
 &= f_Y\left(\frac{\beta z}{2}\right) \left| \frac{d}{dz} \beta \frac{z}{2} \right| \\
 &= \frac{1}{\Gamma(n)\beta^n} \left(\frac{\beta z}{2}\right)^{n-1} e^{-\frac{\beta z}{2}} \left| \frac{\beta}{2} \right| \\
 &= \frac{1}{\Gamma(n)2^n} z^{n-1} e^{-\frac{z}{2}}.
 \end{aligned}$$

So  $Z \sim \text{Gamma}(\beta = 2, \alpha = n)$ , which is a  $\chi^2$  distribution with  $n$  degrees of freedom.

9. The number of patients arriving at the emergency room (ER) at a large city hospital is a discrete random variable that can be modeled with the Poisson distribution with a mean of 10 patients per hour. A random variable  $Y_i$  is defined to represent the amount of time passing between the arrival time of Patient  $i$  and arrival time of Patient  $i - 1$ , where  $i = 2, 3, \dots$ . Specifically,  $Y_i = X_i - X_{i-1}$  where  $X_i$  = arrival time of patient  $i$ ,  $i = 2, 3, \dots$ . A random variable  $Z$  is defined as the total amount of time passing until the fourth patient arrives at the ER in specific one-hour period. What is the probability that at least 10 minutes to at most 20 minutes pass from the start of the hour until the fourth patient arrives at the ER?

Our first thing to notice is that each  $Y_i \sim \text{Exponential}(10)$ , modeling time between events. As well we can say that the MGF of these distributions is equal;  $M_{Y_i}(t) = (1 - \beta t)^{-1}$ .

$Z$  is the total time of the first four events,  $Z = \sum_{i=1}^4 Y_i$ .

From this we can find the MGF of  $Z$ :

$$\begin{aligned}
 M_Z(t) &= E[\exp(tZ)] \\
 &= E[\exp(t \sum_{i=1}^4 Y_i)] \\
 &= E[\prod_{i=1}^4 \exp(tY_i)] \\
 &= \prod_{i=1}^4 E[\exp(tY_i)] && \text{By Independence} \\
 &= M_{Y_i}(t)^4 \\
 &= (1 - 10t)^{-4}.
 \end{aligned}$$

This is the MGF of a gamma distribution, so by theorem 6.1,  $Z \sim \text{Gamma}(\beta = 10, \alpha = 4)$ .

We then can find our probability in R:

```
> pgamma(1/3,4,10)-pgamma(1/6,4,10)
[1] 0.3387469
```

10. Let  $Z$  and  $W$  be two independent random variables, such that  $Z$  follows a standard normal distribution and  $W$  follows a Chi-Squared distribution with  $\nu$  degrees of freedom. Derive the probability density function of the random variable  $T$ , where  $T$  is defined as:  $T = \frac{Z}{\sqrt{\frac{W}{\nu}}}$  ie. find  $f_T(t)$ .

Let  $T = \frac{Z}{\sqrt{\frac{W}{\nu}}}$ . We can also write  $Z = T\sqrt{\frac{W}{\nu}}$ , and  $\frac{Z^2}{T^2}\nu = W$ .

Then since  $W \perp Z$  we can write our joint pdf:

$$f_{W,Z}(w, z) = \frac{\exp\left(-\frac{z^2}{2}\right) w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}.$$

We use the method of transformations:

$$\begin{aligned} f_{T,W}(t, w) &= f_{W,Z}\left(w, t\sqrt{\frac{w}{\nu}}\right) \left| \frac{d}{dt} t\sqrt{\frac{w}{\nu}} \right| \\ &= \frac{\exp\left(-\frac{t^2 w}{2\nu}\right) w^{\frac{\nu+1}{2}-1} e^{-\frac{w}{2}}}{\sqrt{2\pi} 2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}. \end{aligned}$$

And integrate out our variable  $w$ .

$$f_T(t) = \frac{1}{\sqrt{2\pi} \Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty \exp\left(-w\left(\frac{t^2}{\nu} + 1\right)\right) w^{\frac{\nu+1}{2}-1} dw.$$

We recognise a beta distribution which must integrate to 1, provided we extract a factor of:

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\left(\frac{1}{2}\left(\frac{t^2}{\nu} + 1\right)\right)^{(\nu+1)/2}}$$

This gives our final pdf:

$$f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

11. Let  $X_1, X_2, \dots, X_{20}$  represent a random sample, where each  $X_i$  is normally distributed with a mean of  $\mu_X = 100$  and a variance  $\sigma_X^2 = ?$ . From what you know about the distribution of  $(n-1)\frac{s^2}{\sigma^2}$ , find a symmetric interval that will contain  $\sigma_X^2$  95% of the time.

Recall that  $S^2 \frac{n-1}{\sigma^2} \sim \chi_{df=n-1}^2$ .

We first construct a 95% confidence interval for  $\chi_{df=19}^2$  using:

```
> qchisq(.025,19)
[1] 8.906516
> qchisq(.975,19)
[1] 32.85233
```

And then convert this to  $\sigma^2$ .

$$\begin{aligned} 0.95 &= P(8.906516 < \chi_{df=19}^2 < 32.85233) \\ &= P(8.906516 < \frac{19S^2}{\sigma^2} < 32.85233) \\ &= P(0.5783456S^2 < \sigma^2 < 2.13327S^2). \end{aligned}$$

So our interval capturing  $\sigma_X^2$  95% of the time is:

$$(0.5783456 S^2, 2.13327 S^2)$$



12. Referring to Question 11: Suppose that the random sample of  $n = 20$  produced the following data:

92.6, 96.6, 125.4, 115.3, 96.0, 94.6, 102.0, 108.7, 90.6, 101.4,  
91.3, 107.0, 111.9, 98.4, 105.1, 108.2, 96.1, 95.5, 104.7, 100.1

Use your result to find an interval that will 'capture' the value of  $\sigma_X$ , 95% of the time. We simply find the sample variance in R with `var(X)`, which gives us  $S^2 = 78.64724$ .

Then we simply plug this into the interval above, giving us the interval:

$$(45.48529, 167.7758)$$

And since this is our interval for  $\sigma^2$ , we must square root each value to find our interval for  $\sigma$ .

$$(6.744278, 12.95283)$$

13. Let  $S_1^2$  denote the sample variance for a random sample of ten  $\ln(LC50)$  values for copper and let  $S_2^2$  denote the sample variance for a random sample of eight  $\ln(LC50)$  values for lead, both samples using the same species of fish. The population variance for measurements on copper is assumed to be twice the corresponding population variance for measurements on lead. Assume  $S_1^2$  to be independent of  $S_2^2$ .

Find numbers  $a$  and  $b$  such that

$$P\left(a \leq \frac{S_1^2}{S_2^2} \leq b\right) = 0.8.$$

Note the interval should be symmetric.

We begin with:

$$P\left(a \leq \frac{S_1^2}{S_2^2} \leq b\right) = 0.8.$$

And we multiply by  $\frac{\sigma_2}{\sigma_1}$  to get an F-distribution.

$$P\left(\frac{\sigma_2}{\sigma_1}a \leq F_{9,7} \leq \frac{\sigma_2}{\sigma_1}b\right) = 0.8.$$

Assuming that  $\sigma_1 = 2\sigma_2$ , we get

$$P\left(\frac{a}{2} \leq F_{9,7} \leq \frac{b}{2}\right) = 0.8.$$

And in R this gives us:

```
> qf(0.1,9,7)
[1] 0.3991517
```

Which means  $\frac{a}{2} = 0.3991517$  and  $a = 0.7983034$ .

```
> qf(0.9,9,7)
[1] 2.724678
```

And from this,  $\frac{b}{2} = 2.724678$ , meaning  $b = 5.449355$ .

14. Let  $Y_1, Y_2, \dots, Y_5$  be a random sample of size 5 from a normal population with mean 0 and variance 1 and let

$$W = \sum_{i=1}^5 Y_i^2 \text{ and } U = \sum_{i=1}^5 (Y_i - \bar{Y})^2.$$

Before we begin the below questions, let's find the distribution of  $W, U$ .

For  $W$ : By theorem 7.2, The distribution of  $n$  standard normally distributed variables is  $\chi^2$  with  $n$  degrees of freedom. So  $W \sim \chi_{df=5}^2$ .

For  $U$ : Notice  $U = \sum_{i=1}^5 (Y_i - \bar{Y})^2 = \frac{5S^2}{\sigma^2}$ . This is  $\chi_{df=4}^2$ .

What is the distribution of:

- (a)  $\sqrt{5} \frac{Y_6}{\sqrt{W}}$ . Show your work as to how you got this.

Note that  $Y_6 \sim N(0, 1)$ . So we have a standard normal distribution over the root of a chi-square distribution, which has 5 degrees of freedom.

Note that our expression can be rewritten:

$$\sqrt{5} \frac{Y_6}{\sqrt{W}} = \frac{Y_6}{\sqrt{\frac{W}{5}}}.$$

And since  $W$  has  $df = 5$ , we can say that the above expression follows a student's  $t$ -distribution, with 5 degrees of freedom.

- (b)  $\frac{2Y_6}{\sqrt{U}}$ . Show your work as to how you got this.

Rewrite again:

$$\frac{2Y_6}{\sqrt{U}} = \frac{Y_6}{\sqrt{\frac{U}{4}}}.$$

And similarly to above, we can say that the above is a  $t$ -distribution, this time with 4 degrees of freedom.

- (c)  $\frac{2(5\bar{Y}^2 + Y_6^2)}{U}$  Show your work as to how you got this.

Note  $\sqrt{5}\bar{Y} = \frac{\bar{Y} - 0}{\frac{1}{\sqrt{5}}}$ , so  $\sqrt{5}\bar{Y} \sim N(0, 1)$ . Therefore, the square of  $\bar{Y}$  is chi-squared, with 1 degree of freedom.

This gives:

$$\begin{aligned} \frac{2(5\bar{Y}^2 + Y_6^2)}{U} &= \frac{4(\chi_{df=1}^2 + \chi_{df=1}^2)}{2\chi_{df=4}^2} \\ &= \frac{\frac{\chi_{df=2}^2}{2}}{\frac{\chi_{df=4}^2}{4}} \sim F_{2,4} \end{aligned}$$

.