

Math 151 - Probability Theory - Homework 8

your name here

Due Friday, March 15, 2019

```
knitr::opts_chunk$set(message=FALSE, warning=FALSE, fig.height=3, fig.width=5,
  fig.align = "center")
```

[7] DeGroot, section 3.8

Suppose that the random variable X has the uniform distribution on $[0, 1]$. Determine the p.d.f. of

- (a) X^2
- (b) $-X^3$
- (c) $X^{1/2}$

[9] DeGroot, section 3.8

Suppose that X has the uniform distribution on the interval $[0, 1]$. Construct a random variable $Y = r(X)$ for which the p.d.f. will be

$$g(y) = \begin{cases} \frac{3}{8}y^2 & \text{for } 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

[4 & 5] DeGroot, section 3.9 Hint: draw a picture for each of these parts! Really, I mean it, **draw pictures**.

Suppose that X_1 and X_2 have a continuous joint distribution for which the joint p.d.f. is as follows:

$$f(x_1, x_2) = \begin{cases} x_1 + x_2 & \text{for } 0 < x_1 < 1 \text{ and } 0 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

[4] Find the p.d.f. of $Y = X_1X_2$.

[5] Find the p.d.f. of $Z = X_1/X_2$.

[14] DeGroot, section 3.9

Suppose that X_1, \dots, X_n form a random sample of n observations from the uniform distribution on the interval $[0, 1]$, and let Y denote the second largest of the observations. Determine the pdf of Y . Hint: first determine the cdf of Y by noting that:

$$F_Y(y) = P(Y \leq y) = P(\text{at least } n - 1 \text{ observations } \leq y).$$

[R1] We are designing a standardized test, and we want to have some sense of how variable the test is. We want to use the range of scores to estimate the variability.

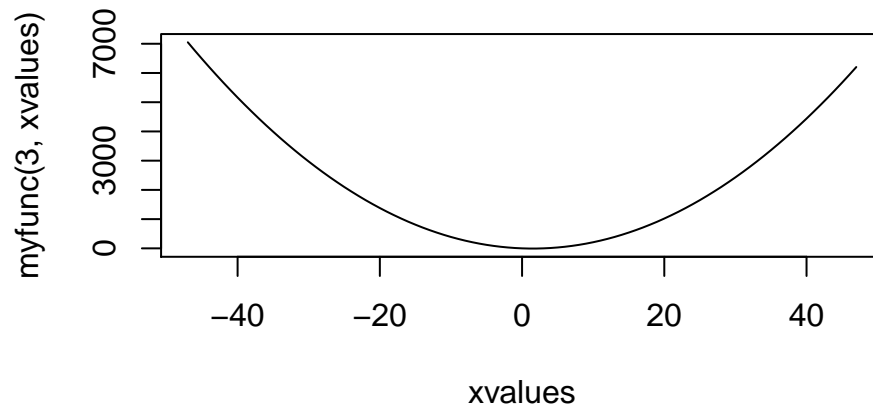
Consider a sample of size 100 from a $U[0, 1]$ distribution.

- (a) Find the pdf of $Y = X_{\max} - X_{\min}$. Plot the pdf (as a line function). (Hint: see your text and examples 3.9.7 and 3.9.8.)

```
# Here is some code that let's me plot a function of interest
```

```
myfunc <- function(a,b) { a*b^2 - b*a^2 }
```

```
xvalues <- seq(-47,47,1) # figure out what this line of code does!
plot(xvalues, myfunc(3, xvalues), type="l") # try values for `a` other than 3
```



also, p.s. type="l" is for `line`, it is not the number one

- (b) Find the pdf (using a histogram) of $Y = X_{\max} - X_{\min}$ using simulations. Plot the pdf (using a histogram, not a line). Make sure your code is reproducible.

Hint: your R code should have a loop which does something 1000 or so times. Each time you go through the loop, create a sample of uniform values, find the max, find the min, subtract them, keep the value. After you've gone through the loop 1000 times you should have 1000 differences, make a histogram. Look at: `?runif`.

Ask me / Piazza if you aren't sure what the R code should look like!!!

- (c) Change 1: Suppose we are concerned with outliers, so we want to use the IQR, ($Y = X_{(75)} - X_{(25)}$). That is, our variable of interest is the 75 percentile minus the 25 percentile.

Comment on the difficulties of the analytic solution for this problem (that is: why is it now much harder to find the analytic solution?), and perform the R computation to find the distribution of the IQR. Plot the pdf of the IQR (a histogram). How does the distribution of the IQR change as compared to the distribution of the max - min?

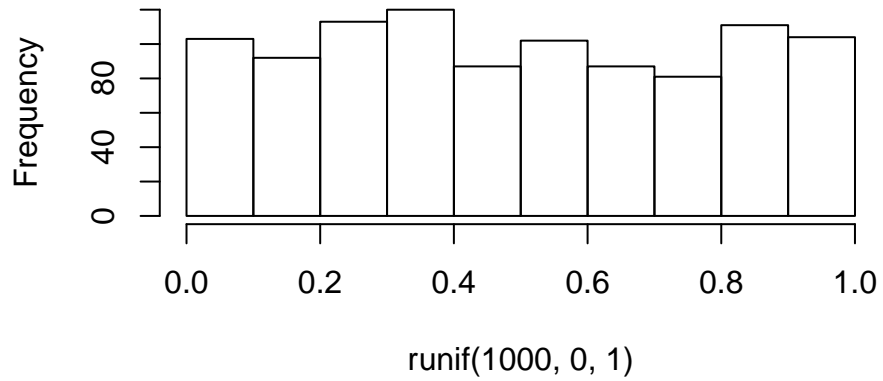
- (d) Change 2 (back to max-min): Now suppose that the scores are actually normally distributed.

Comment on the difficulties of the analytic solution (for max-min) (why is it much harder to find the analytic solution now?) and perform the R computation using normally distributed data (mean 0.5, sd = 0.2). Plot the pdf of the max-min for normal data (a histogram). How does the distribution of the max-min change when the data are normally distributed? [n.b. I set the normal parameters to give a distribution which is pretty similar to the $U[0,1]$ set up. To test that try: `hist(rnorm(1000, mean = 0.5, sd = 0.2))`.]

#Not part of the HW for you to turn in, but just for you to see:

```
hist(runif(1000,0,1))
```

Histogram of runif(1000, 0, 1)



```
hist(rnorm(1000,mean=.5, sd=.2))
```

Histogram of rnorm(1000, mean = 0.5, sd = 0.2)

