Math 152 - Statistical Theory - Homework 1

Ethan Ashby

Due: Friday, August 28, 5pm PDT

Assignment

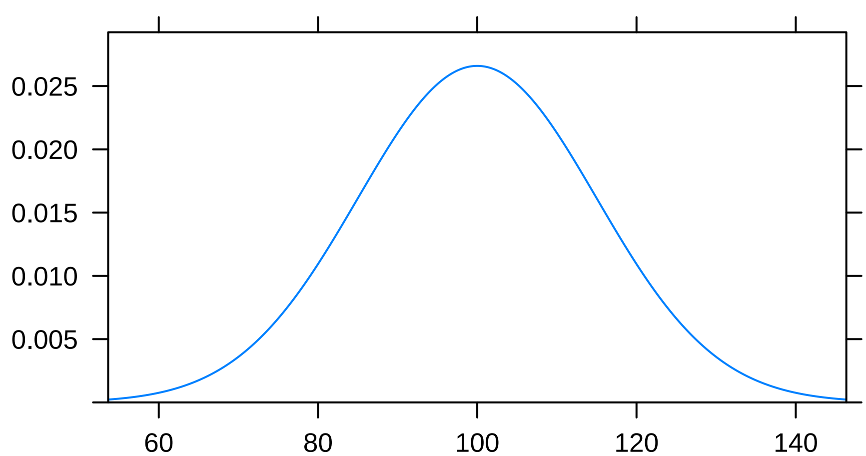
You might be interested in these resources:

- Probability distribution demos: ISLE widgets
- Fantastic cheat sheet for Univariate Distributions created by Lindsey Leemis

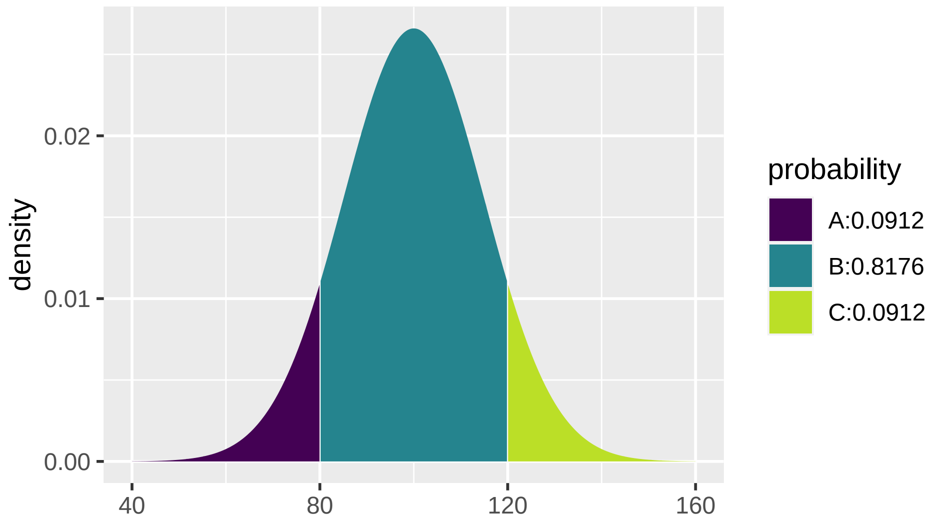
1: PodQ Describe one thing you learned from someone in your pod this week (it could be: content, logistical help, background material, R information, etc.) 1-3 sentences.

2: R - code provided, use help() Children's IQ scores are normally distributed with a mean of 100 and a standard deviation of 15. One might, for example, be interested in the proportion of children having an IQ between 80 and 120. One might also randomly generate 47 students from such a population and create a histogram.

```
set.seed(47)  # to set the randomness

plotDist('norm', mean = 100, sd = 15)
```

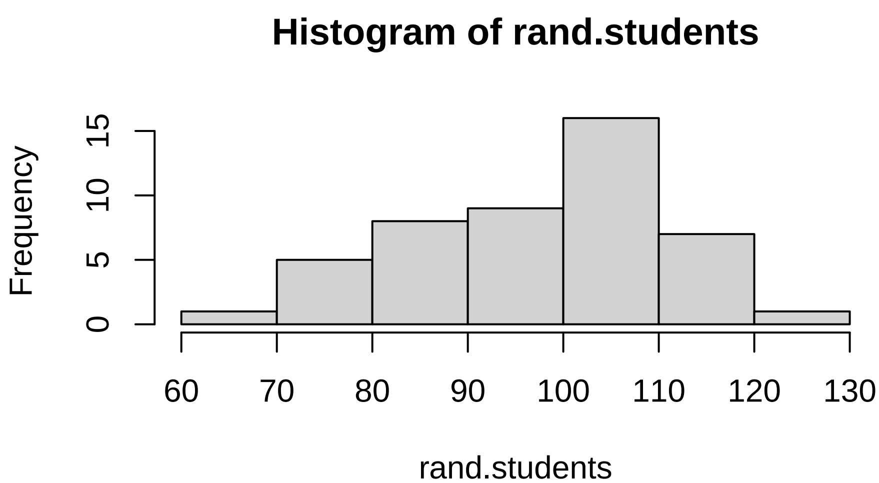



```
xpnorm(c(80, 120), mean = 100, sd = 15)
```




```
## [1] 0.09121122 0.90878878
```

```r
rand.students = rnorm(47, mean = 100, sd = 15)
hist(rand.students)
```




```r
knitr::kable(data.frame("Mean"=mean(rand.students), "Variance"=var(rand.students), "Standard Deviation"=
```

Mean	Variance	Standard.Deviation
97.34426	182.326	13.50282

a. The R function `dnorm()` gives the height of the density (i.e., f(x), the pdf). What is a density?

I consulted my Math151 notes on this section

The density or probability density function (pdf), is a function of a continous random variable X to $\mathbb{R}$. If the probability that the random variable, X lies in the range $[A, B]$ is defined as:

$$P(A \leq X \leq B) = \int_a^b f(x)dx$$

then it is a p.d.f.

This p.d.f is endowed with a number of important properties.

1. The density is nonegative for all values of $x\ldots$ it can be $0\ldots$ can be greater than 1 too
2. The density integrates to 1 over all value of x
3. The density need not be continuous

b. The R function `pnorm()` gives the area associated with the density. What is the area associated with the density? [Note that the expanded function `xpnorm()` which is available in the `mosaic` package includes pictures with the associated areas.]

The area under the density curve over some interval $[A, B]$ corresponds to the probability of your random variable X landing somehwere in that interval. This probability is non-negative and with least upper bound 1. `pnorm(q, mean, sd)` (when the `lower.tail=TRUE`) gives us the probability that x lies in the interval $[-\infty, q]$ when $X \sim \mathcal{N}(mean, sd)$.

c. Look up the R function `qnorm()`. What does `qnorm()` provide? [Note that the expanded function `xqnorm()` which is available in the `mosaic` package includes pictures with the associated quantiles.]

`qnorm(x, mean, sd)` yields the quantile associated with a value of x on a normal distribution with `mean=mean` and `sd=sd`. For example, `qnorm(0)` yields 0.5, meaning that the 0 is the median of the of a normal distribution with mean=0 and sd=1.

d. Look up the R function `rnorm()`. What does `rnorm()` provide?

`rnorm(x, mean, sd)` generates `x` random values from a normal distribution with mean=`mean` and sd=`sd`.

3: R - simulating data Consider the beta distribution (Google it if you are unfamiliar with it).

a. Provide the pdf for any *symmetric* beta distribution (specify a and b which should not equal to 1).

- What are the theoretical mean, variance, and standard deviation of your beta distribution? (not found using R)
- Using the R function `rbeta()`, generate 100 random beta variates, and make a histogram of the random sample.
- Find the sample mean, sample variance, and sample standard deviation of your 100 random deviates.
- plot the density (pdf); use `plotDist()`

The pdf for a beta distribution with shape parameters a, b is symmetric when $a = b$.

$$f(x) = \begin{cases} \frac{x^{(a-1)}(1-x)^{(b-1)}}{B(a,b)} & 0 < x < 1 \\ 0 & else \end{cases}$$

Where $B(a, b) = \int_0^1 x^{(a-1)}(1-x)^{(b-1)}dx$.

I choose $a, b = 0.5$. So I get:

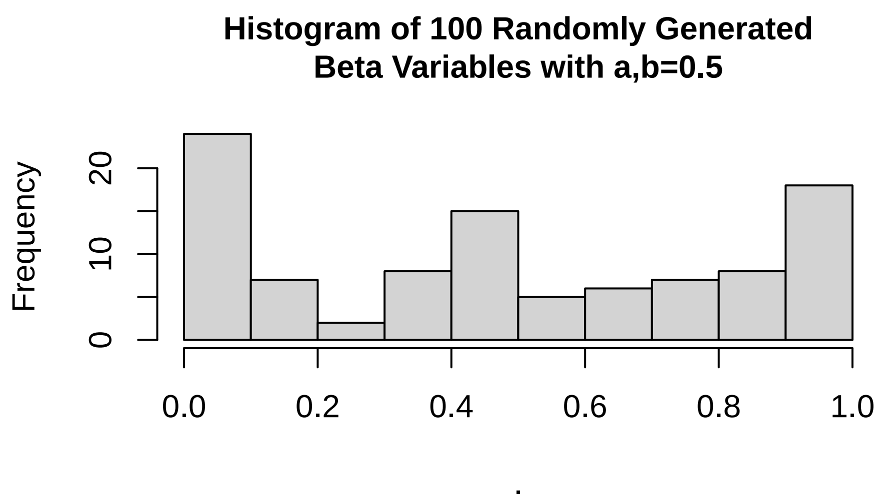
$$f(x) = \frac{x^{-0.5}(1-x)^{-0.5}}{\int_0^1 x^{-0.5}(1-x)^{-0.5}dt}$$

The **mean** $E[X] = \frac{a}{a+b} = 0.5$

The **variance** $var[X] = \frac{ab}{(a+b)^2(a+b+1)} = \frac{1}{8}$

The **standard deviation** is the square root of the variance: $\sqrt{var[X]} = \frac{1}{2\sqrt{2}}$

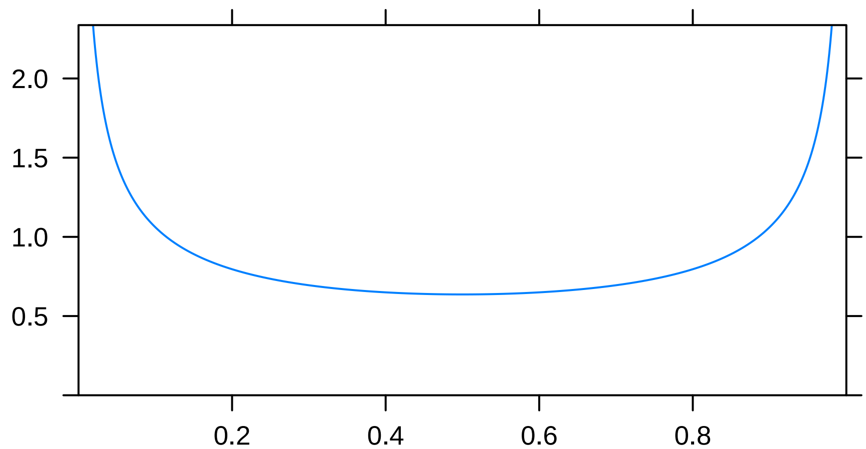
```
set.seed(47)
sample<-rbeta(100, 0.5, 0.5)

#histogram
sample %>% hist(main="Histogram of 100 Randomly Generated\nBeta Variables with a,b=0.5", cex.main=1)
```




```
#summary
data.frame(data=sample) %>% summarize(mean=mean(data), variance=var(data), sd=sd(data))
```

```
##        mean  variance        sd
## 1 0.4785855 0.1223126 0.3497322
```

```
#plotDist of this beautiful symmetric distribution!
plotDist('beta', params=c(0.5, 0.5))
```



b. Provide the pdf for any **non-symmetric** beta distribution (specify a and b).

- What are the theoretical mean, variance, and standard deviation of the beta distribution? (not found using R)
- Using the R function `rbeta()`, generate 100 random beta variates, and make a histogram.
- Find the sample mean, sample variance, and sample standard deviation of your 100 random deviates.
- plot the density (pdf)

To get a **non-symmetric** beta distribution, we need a and b to be different values.

I choose $a = 0.5, b = 1.5$. So we get:

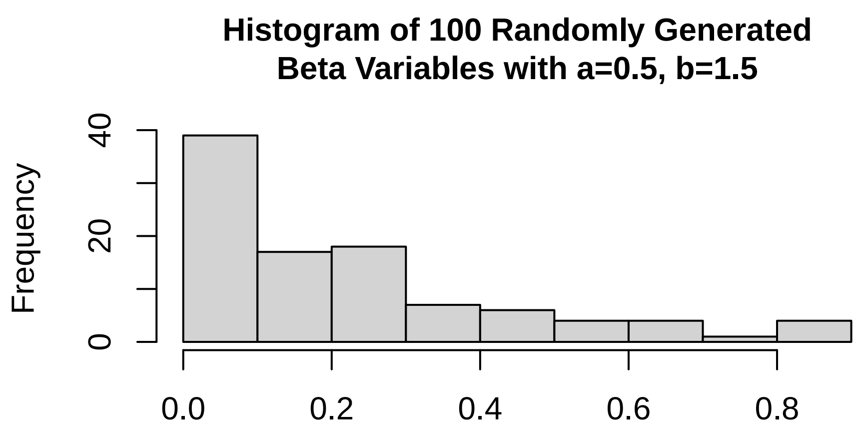
$$f(x) = \frac{x^{-0.5}(1-x)^{0.5}}{\int_0^1 x^{-0.5}(1-x)^{0.5}dt}$$

The **mean** $E[X] = \frac{a}{a+b} = 0.25$

The **variance** $var[X] = \frac{ab}{(a+b)^2(a+b+1)} = \frac{1}{16}$

The **standard deviation** is the square root of the variance: $\sqrt{var[X]} = \frac{1}{4}$

```r
set.seed(47)
sample<-rbeta(100, 0.5, 1.5)

#histogram
sample %>% hist(main="Histogram of 100 Randomly Generated\nBeta Variables with a=0.5, b=1.5", cex.main=
```

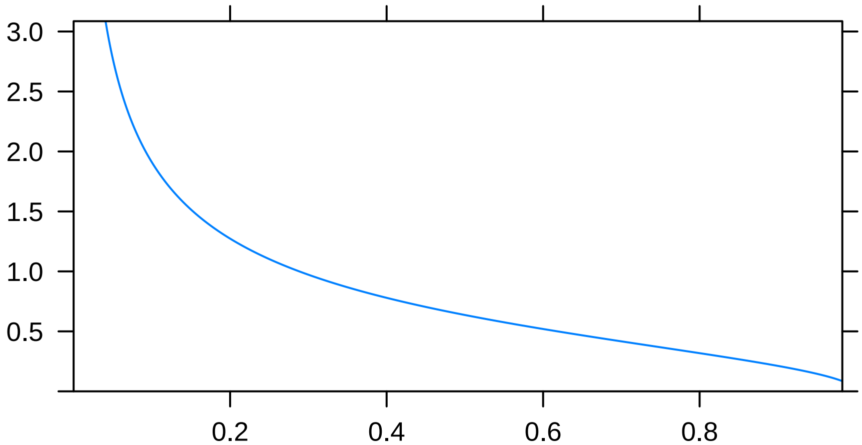


.

```r
#summary
data.frame(data=sample) %>% summarize(mean=mean(data), variance=var(data), sd=sd(data))
```

```
##        mean   variance        sd
## 1 0.2168337 0.05121106 0.2262986
```

```r
#plotDist of this beautiful symmetric distribution!
plotDist('beta', params=c(0.5, 1.5))
```


3.0
2.5
2.0
1.5
1.0
0.5
0.2
0.4
0.6
0.8