Analytics 511: Homework #2

Brody Vogel

September 16th

Preparation

```r
set.seed(1234)
```

Problem 9

$Pr(X \leq 10) \approx .4506$

$Pr(X > 5) = 1 - Pr(X \leq 5) \approx .8492$

$Pr(|X - 8| < 3) = Pr(X < 11) - Pr(X \leq 5) \approx .3558$

$z \approx 4.0258$

```r
pgamma(10, shape = 2.5, scale = 5)
```

```
## [1] 0.450584
```

```r
1 - pgamma(5, shape = 2.5, scale = 5)
```

```
## [1] 0.849145
```

```r
pgamma(10.9999, shape = 2.5, scale = 5) - pgamma(5, shape = 2.5, scale = 5)
```

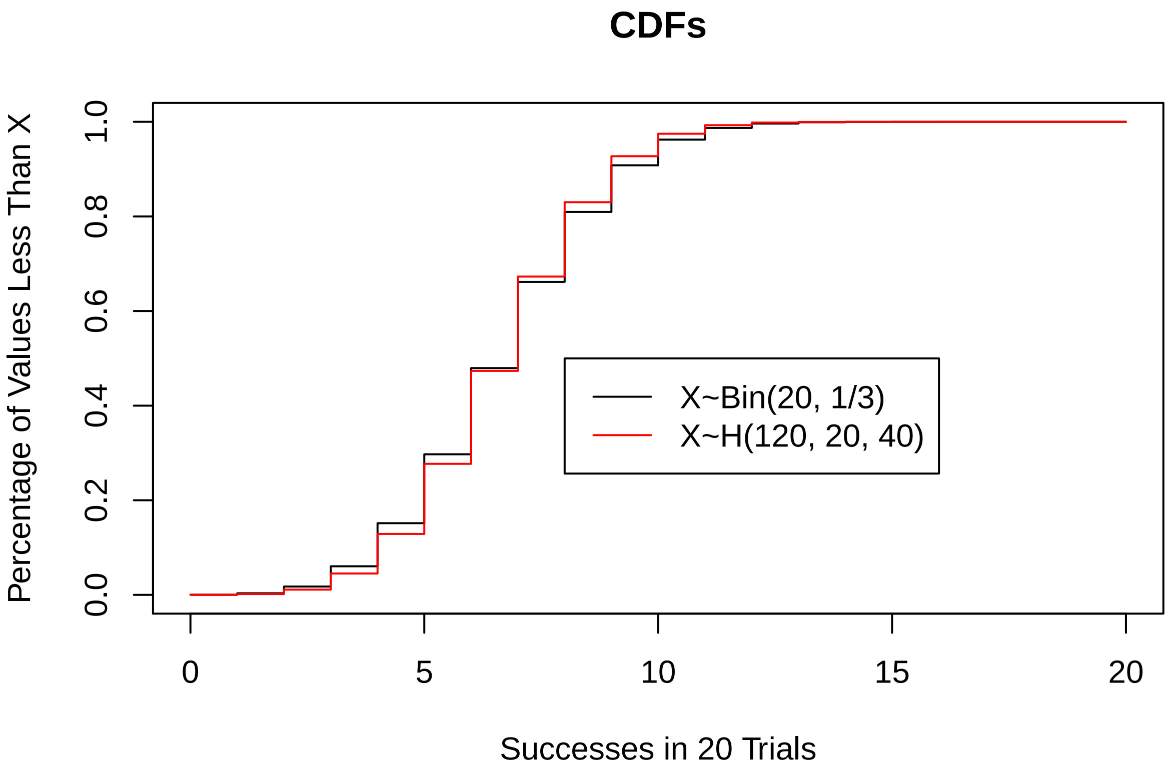
```
## [1] 0.3557661
```

```r
z <- qgamma(.1, shape = 2.5, scale = 5)

z
```

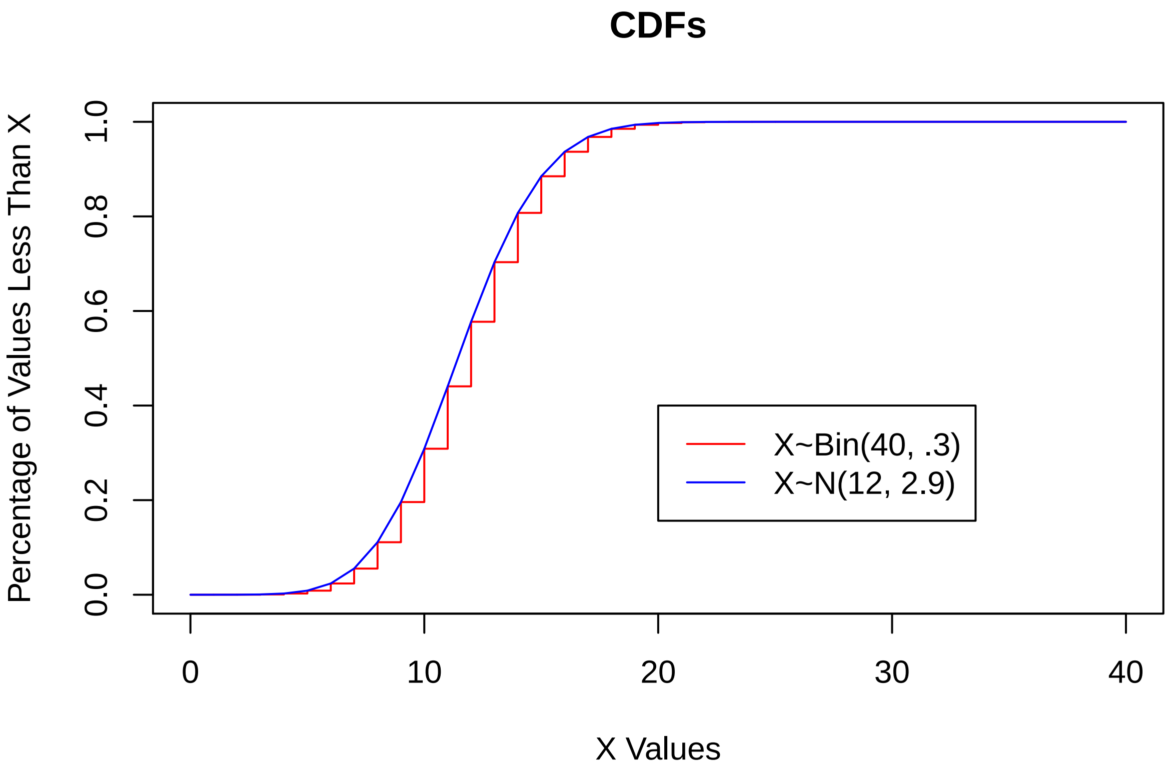
```
## [1] 4.02577
```

Problem 10



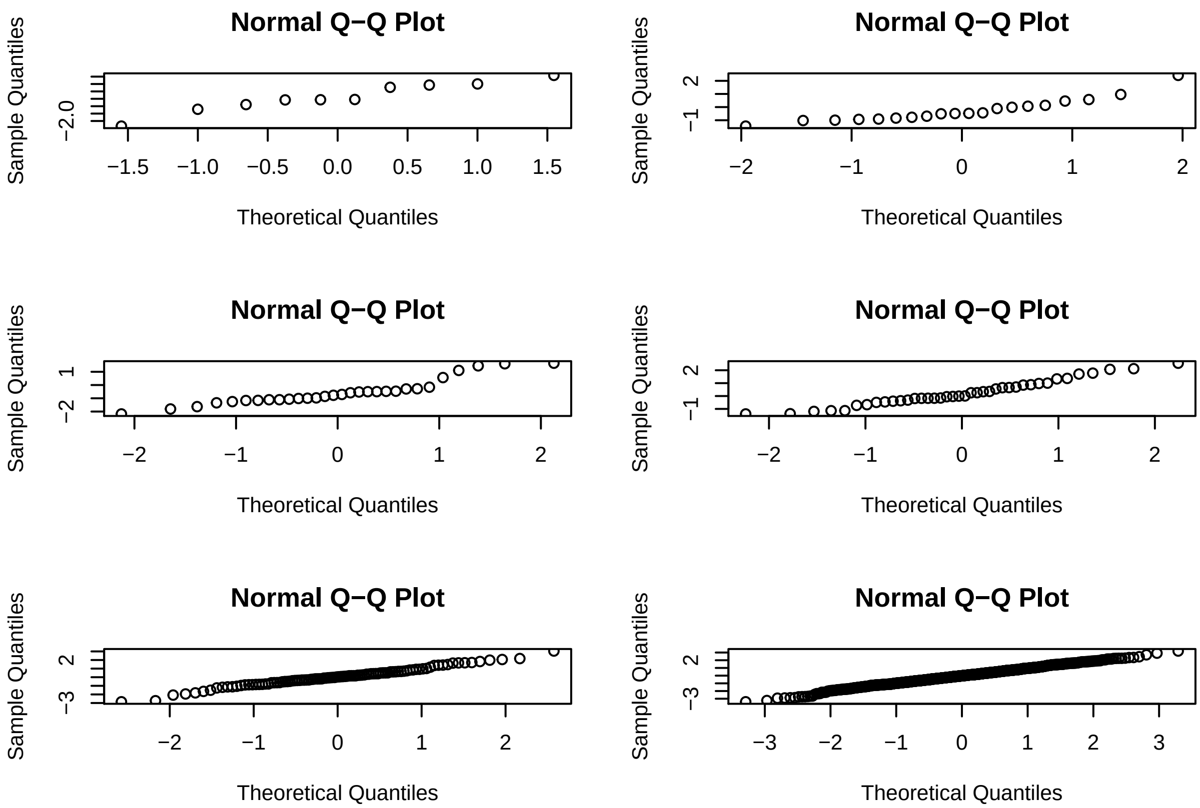
The plot confirms that the distributions are nearly identical.

Problem 11



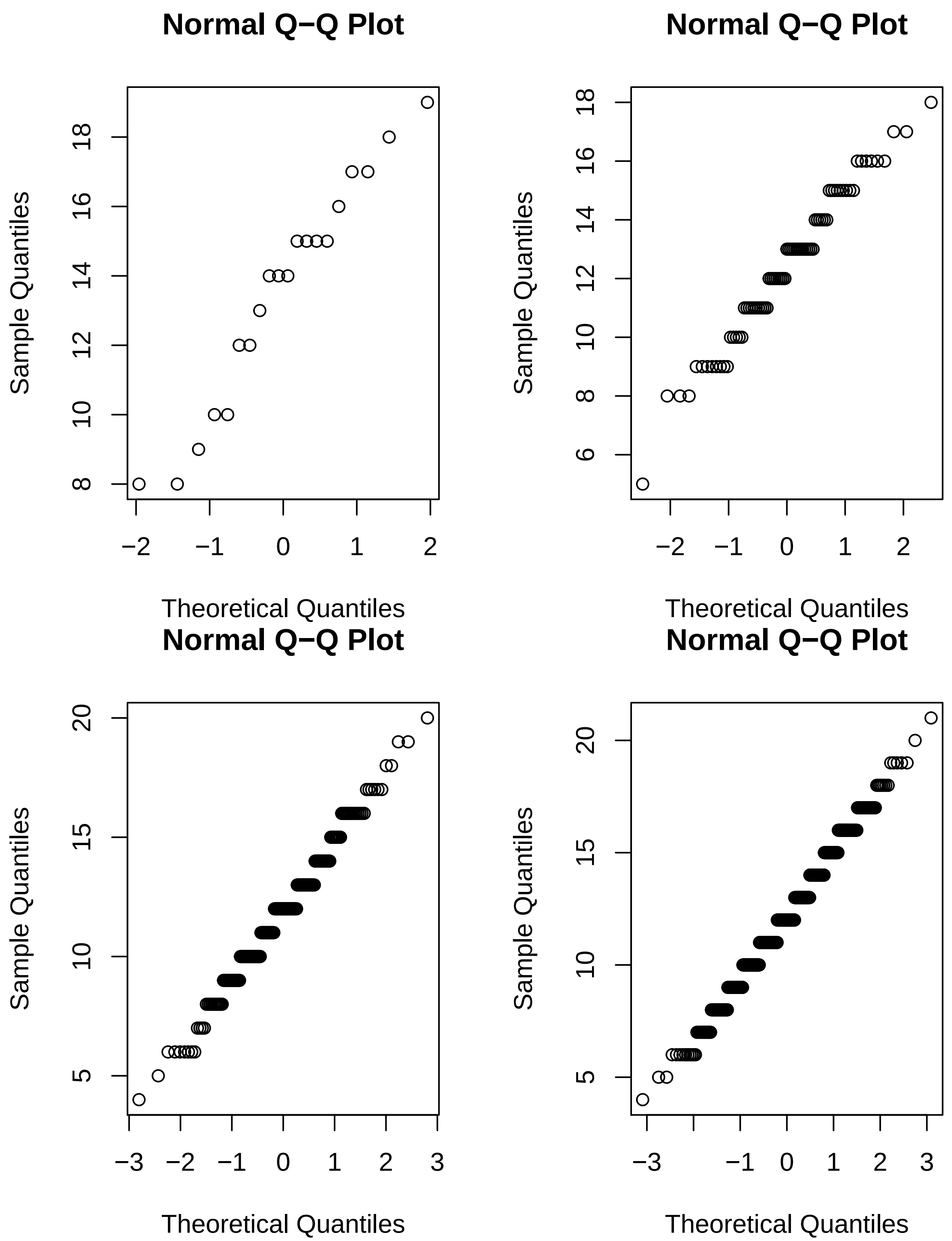
The two distributions are very close, so much so that if you plot them using the same linetype they overlap. The biggest difference is that the Normally-distributed CDF is continuous, whereas the Binomial is discrete. Also, and relatedly, the Binomially-distributed line hits 1.0 exactly at 40, whereas the Normally-distributed line never reaches 1 because it has an infinite upper bound.

Problem 12

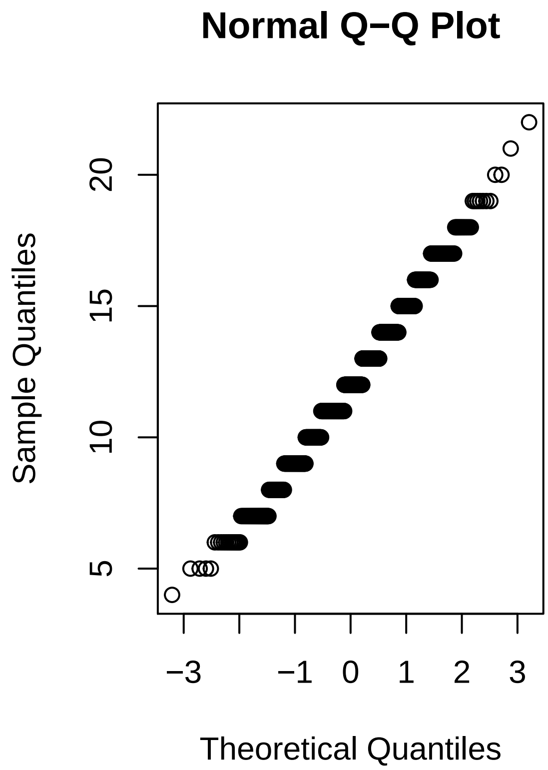


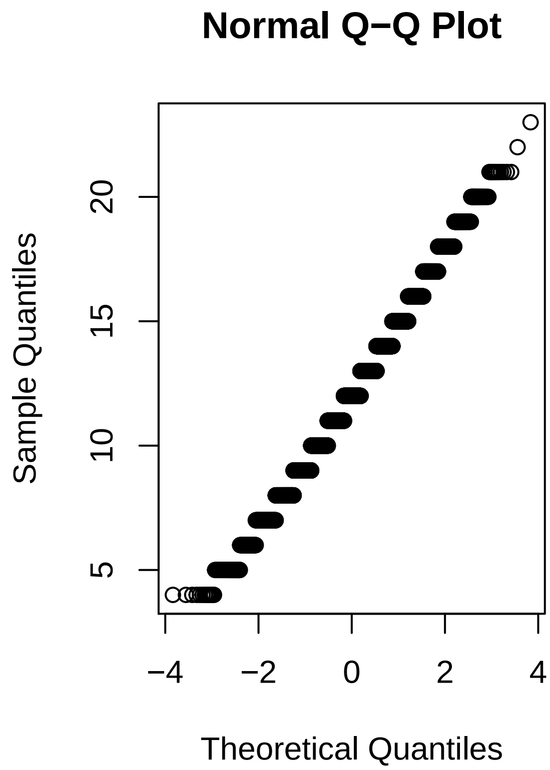
The plots get very close to a straight line as the number of random normal observations is increased. The furthest deviations from a sraight line occur at the ends of the plots where outliers can change the shape.

Problem 13



Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles





The plots come out looking like staircase plots because the samples are from a discrete distribution. So, instead of looking like a continuous line, the steps get longer as the sample size increases, because there are more observations of discrete values. Hence, the rbinom(x, 40, .3) call, for example, will never produce the value 12.5, because it's impossible to get 12.5 successes in 40 trials; it will, however, produce longer steps as x increases.

Problem 14

a) 1 - Pr(X $\leq$ 3) $\approx$ .00135

b) 1 - Pr(X $\leq$ 42) $\approx$ .12167

c) Pr(k = 10) = $\binom{10}{10}.8^{10} * .2^{0} \approx$ .10737

d) F(.9) = .9/1 = .9

e) 1 - F(6.5) $\approx$ .03877

```
a <- 1 - pnorm(3)
b <- 1 - pnorm(42, mean = 35, sd = 6)
c <- .8^10
d <- (.9 - 0) / 1
e <- 1 - pchisq(6.5, df = 2)

a
```

```
## [1] 0.001349898
```

```
b
```

```
## [1] 0.1216725
```

```
c
```

```
## [1] 0.1073742
```

```
d
```

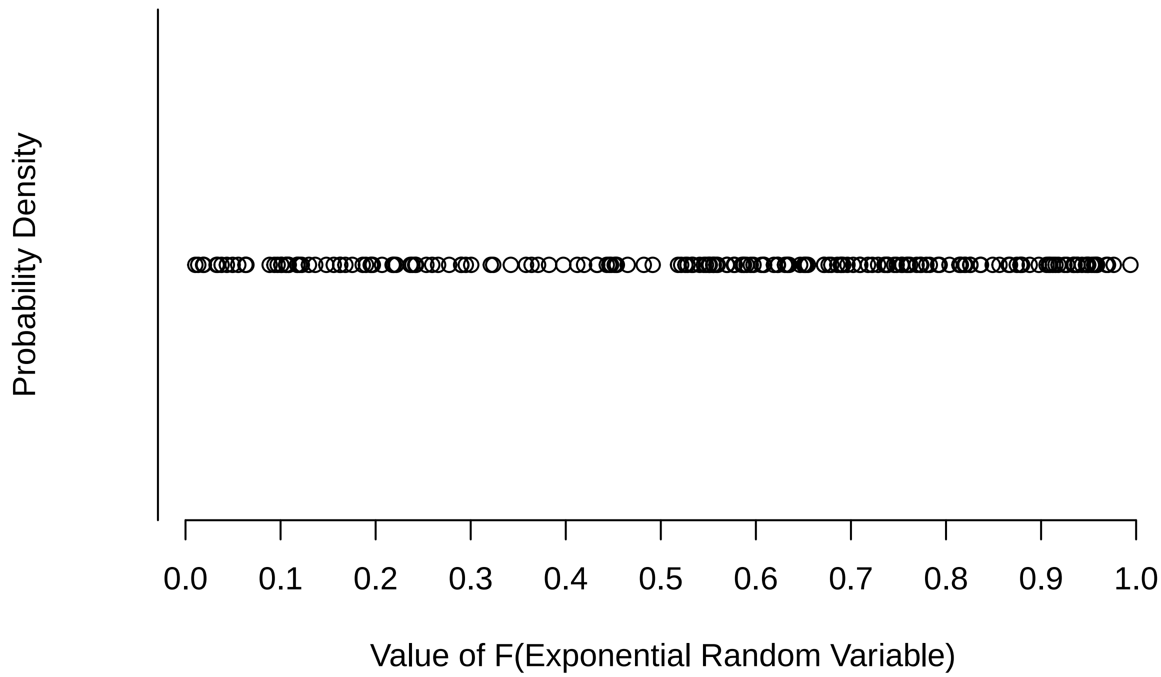
```
## [1] 0.9
```

```
e
```

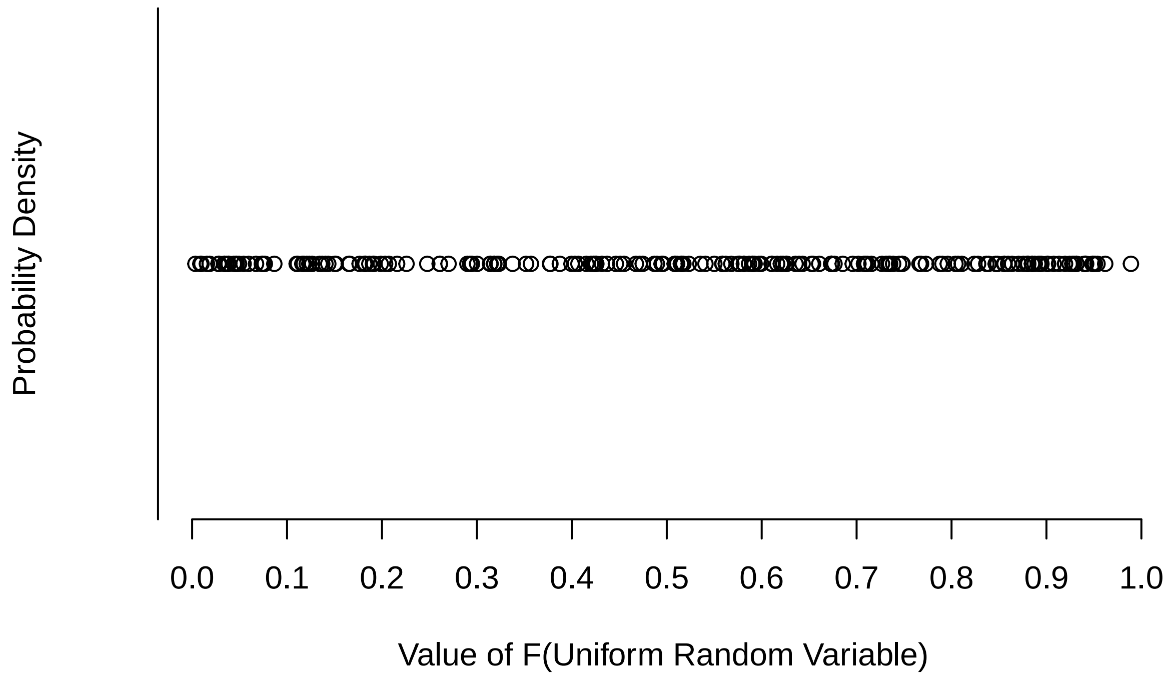
```
## [1] 0.03877421
```

Problem 15

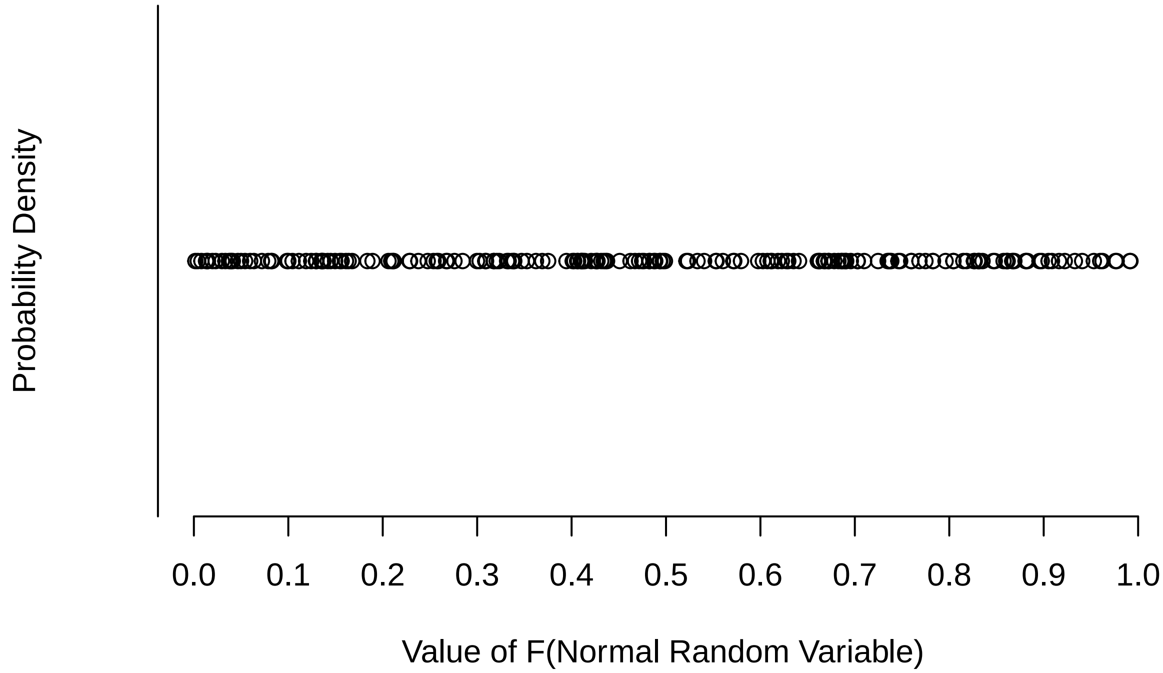
```
options(digits = 20)
exps <- rexp(200)
exps <- pexp(exps)
exps1 <- table(exps)/200
plot(exps, exps1, axes = FALSE, ylab = 'Probability Density', xlab = 'Value of F(Exponential Random Var
axis(1, at = seq(0, 1, .1))
axis(2, at = seq(-.2, 1, .05))
```




```
unis <- runif(200)
unis <- punif(unis)
unis1 <- table(unis) / 200
plot(unis, unis1, axes = FALSE, ylab = 'Probability Density', xlab = 'Value of F(Uniform Random Variabl
axis(1, at = seq(0, 1, .1))
axis(2, at = seq(0, 1, .2))
```




```
norms <- rnorm(200)
norms <- pnorm(norms)
norms1 <- table(norms) / 200
plot(norms, norms1, axes = FALSE, ylab = 'Probability Density', xlab = 'Value of F(Normal Random Variabl
axis(1, at = seq(0, 1, .1))
axis(2, at = seq(-1, 1, .1))
```



Each new random variable, U = F(X), appears to be ~Unif(0,1), which can be proven as follows.

-Let X~EXP(λ). -Because X is continuous, no two values from any sample will be equal. -Because no two values of X from any sample will be equal, no two values of F(X) will be equal. -So our new random variable, U = F(X), is uniformly distributed at 1 / size of the sample. -Also, because $\Pr(X = n, n \in R) = 0$, the point at which U is uniformly distributed will approach 0 as the sample size approaches infinity.

Problem 16

$\alpha \approx .26$

```
expys <- replicate(2000, rexp(1) + rexp(1))
qqnorm(expys^.26)
qqline(expys^.26)
```