```r
library(ggplot2)
library(lattice)
```

Problems 2, 5, 12, 13, 14, 15, 17, 21

Problem 2

Consider the population {3, 6, 7, 9, 11, 14}. For samples of size 3 without replacement, find (and plot) the sampling distribution of the minimum. What is the mean of the sampling distribution? The statistic is an estimate of some parameter—what is the value of that parameter?

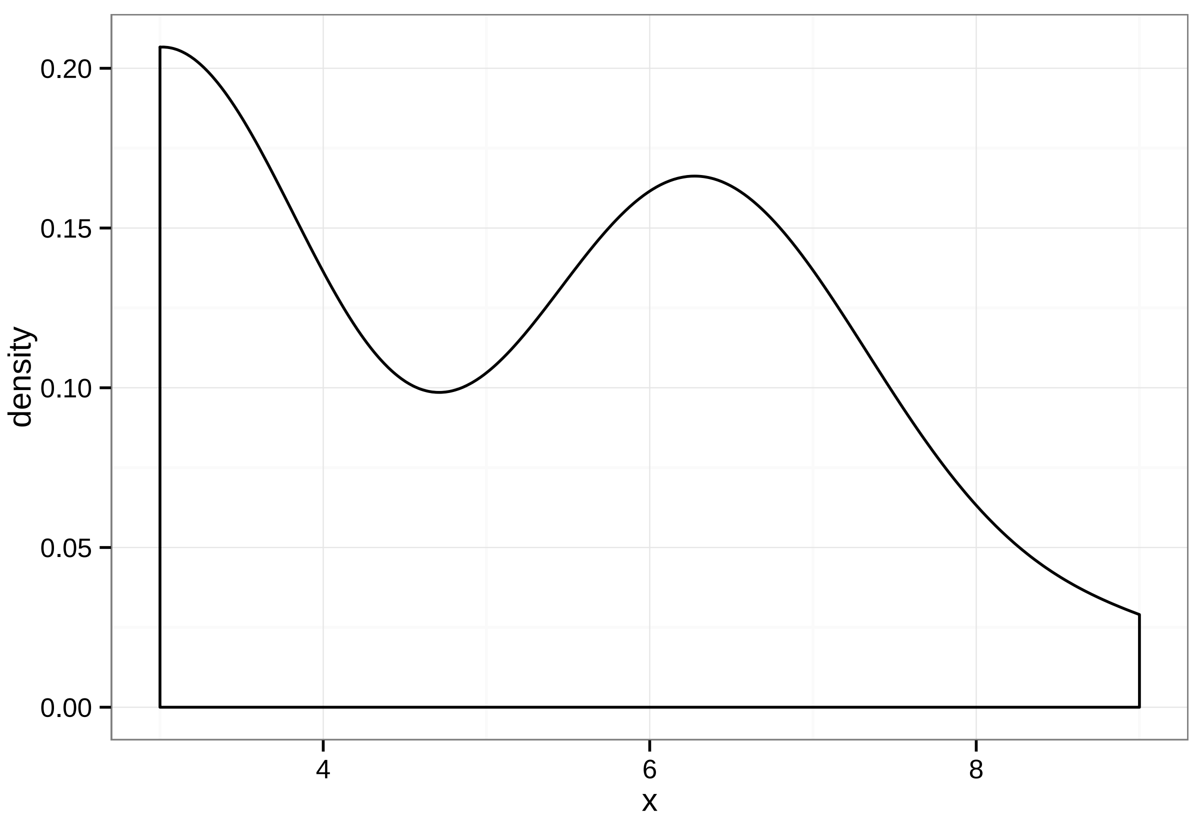
```r
Population <- c(3, 6, 7, 9, 11, 14)
MIN <- apply(combn(Population, 3), 2, min)
MIN
```

```
##  [1] 3 3 3 3 3 3 3 3 3 3 6 6 6 6 6 6 7 7 7 9
```

```r
mean(MIN)  # mean of the sampling distribution
```

```
## [1] 4.8
```

```r
# This statistic is an estimate of the population Min (3).
ggplot(data = data.frame(x = MIN), aes(x = x)) + geom_density() + theme_bw()
```



Prob 2 answer: The mean of the sampling distribution is **4.8** and the value of the parameter is **3**.

Problem 5

Let $X_1, X_2, \ldots, X_n$ be a random sample from some distribution and suppose $Y = T(X_1, X_2, \ldots, X_n)$ is a statistic.

Suppose the sampling distribution of Y has pdf $f(y) = \frac{3}{8}y^2, 0 \le y \le 2$. Find $P(0 \le Y \le \frac{1}{5})$.

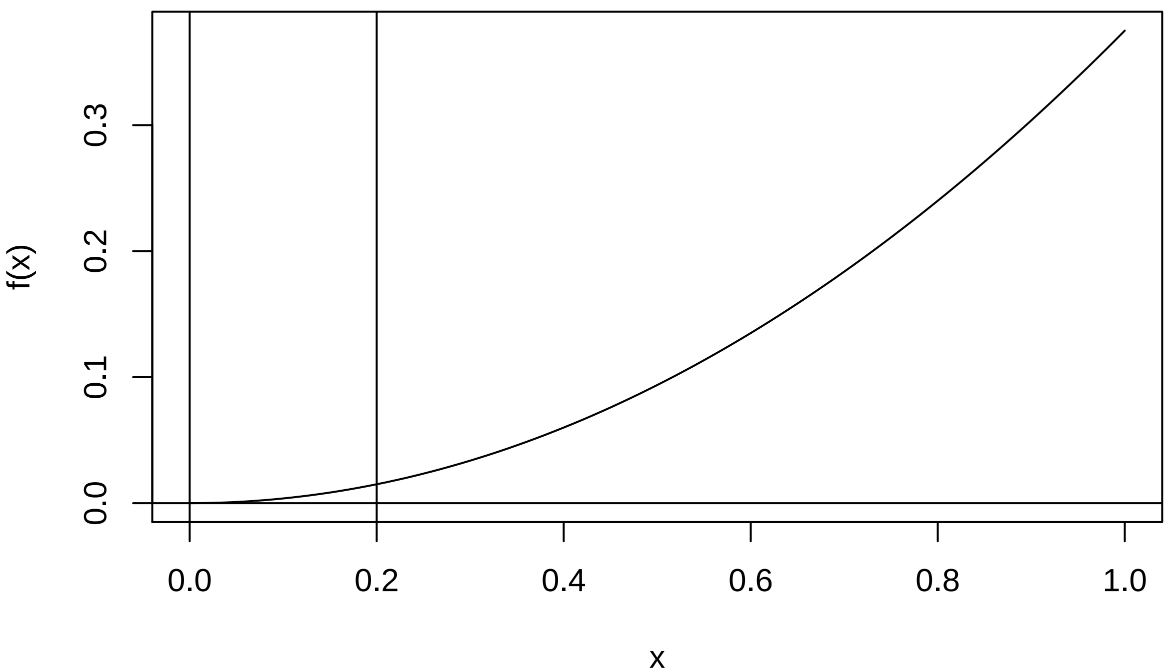
```
f <- function(x){3*x^2/8}
integrate(f, 0, 2)$value # must be 1 for it to be a valid pdf
```

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

```
integrate(f, 0, 1/5)$value
```

```
## [1] 0.001
```

```
curve(f)
abline(v=0)
abline(v=1/5)
abline(h=0)
```



Prob 5 answer: The probability that 0 <= Y <= 1/5 is **0.001**.

Problem 12

A friend claims that she has drawn a random sample of size 30 from the exponential distribution with $\lambda = 1/10$. The mean of her sample is 12.

(a) What is the expected value of a sample mean?
(b) Run a simulation by drawing 1000 random samples, each of size 30, from Exp(1/10) and then compute the mean. What proportion of the sample means are as large as or larger than 12?
(c) Is a mean of 12 unusual for a sample of size 30 from Exp(1/10)?

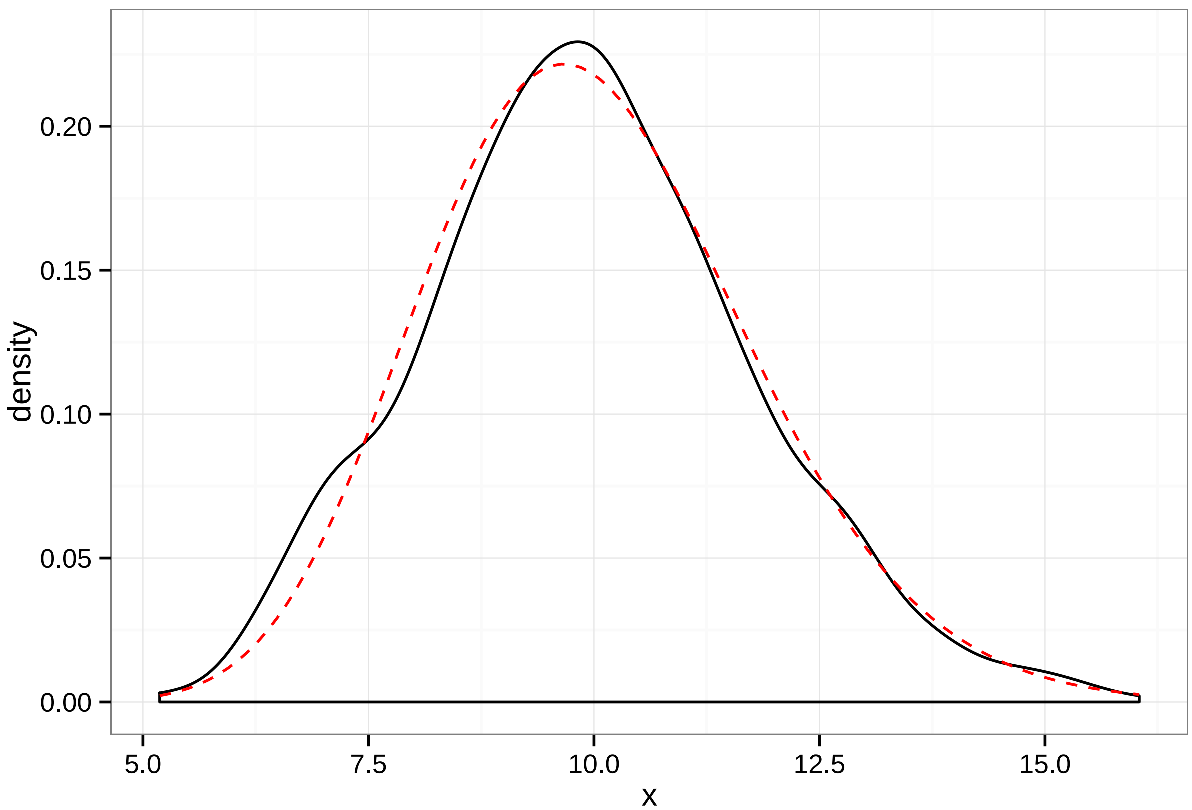
```
set.seed(13)
sims <- 1000
xbar <- numeric(sims)
for(i in 1:sims){
  xbar[i] <- mean(rexp(30, 1/10))
}
mean(xbar)
```

```
## [1] 9.934
```

```
mean(xbar >= 12)
```

```
## [1] 0.126
```

```
library(ggplot2)
ggplot(data = data.frame(x = xbar), aes(x = x)) + geom_density() + theme_bw() + stat_function(fun = dgam
```



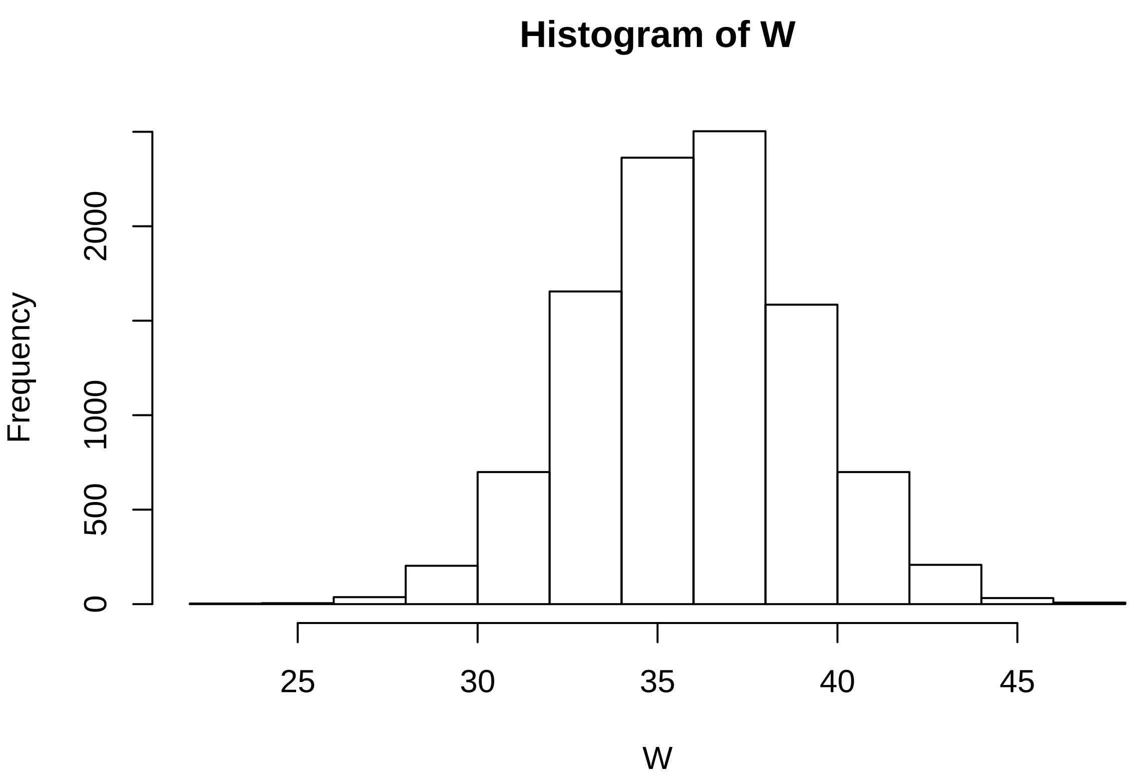
Prob 12 answer: The expected value of a sample mean is $1/\lambda$, in this case **10**. The value of the sample mean is **9.9335**. The proportion of sample means as large as or larger than 12 is **0.126**. Because **0.126** is greater than 0.05, we cannot say that a mean of 12 is unusual for a sample of size 30 from Exp(1/10).

Problem 13

Let $X_1, X_2, \ldots X_{10} \overset{iid}{\sim}$ N(20, 82) and $Y_1, Y_2, \ldots Y_{15} \overset{iid}{\sim}$ N(16, 72). Let W $= \hat{X} + \hat{Y}$.

(a) Give the exact sampling distribution of W.
(b) Simulate the sampling distribution in R and plot your results. Check that the simulated mean and standard error are close to the theoretical mean and the standard error.
(c) Use your simulation to find P(W $<$ 40). Calculate an exact answer and compare.

```
set.seed(13)
sims <- 10000
xbar <- numeric(sims)
ybar <- numeric(sims)
for(i in 1:sims){
  xbar[i] <- mean(rnorm(10, 20, 8))
  ybar[i] <- mean(rnorm(15, 16, 7))
}
W <- xbar + ybar
hist(W)
```



```r
mean(W)  # close to 36
```

```
## [1] 36
```

```r
sd(W)    # close to 3.11
```

```
## [1] 3.055
```

```r
mean(W < 40)
```

```
## [1] 0.9053
```

```r
foo <- sqrt(8^2/10 + 7^2/15)
# Exact answer
pnorm(40, 36, foo)
```

```
## [1] 0.9009
```

Prob 13(a) answer: The exact sampling distribution of W = $\hat{X}$ + $\hat{Y}$ ~ **N(35.9957,9.3344)**.

Prob 13(b) answer: See above.

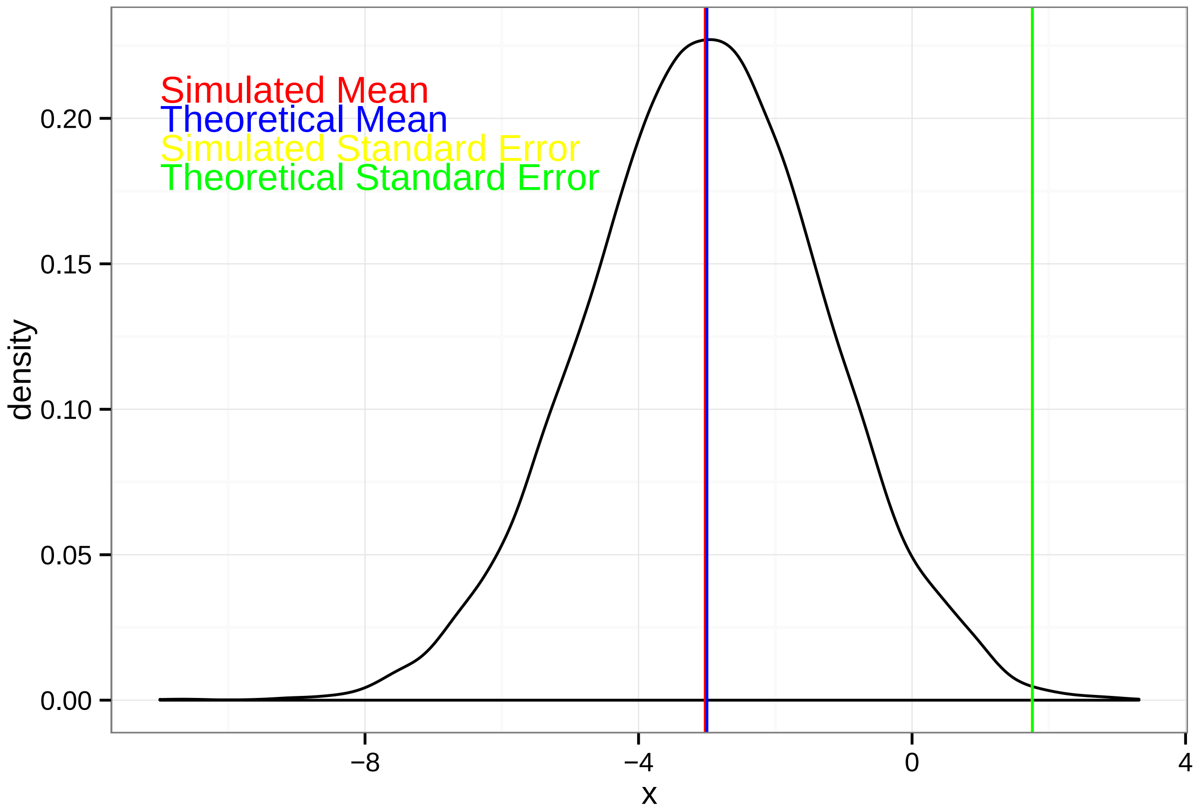
Prob 13(c) answer: From our simulation, P(W < 40) yields **0.9009**.

Problem 14

Let $X_1, X_2, \ldots, X_9 \overset{iid}{\sim} N(7, 3^2)$ and $Y_1, Y_2, \ldots, Y_{12} \overset{iid}{\sim} N(10, 5^2)$. Let $W = \hat{X} - \hat{Y}$.

(a) Give the exact sampling distribution of W.
(b) Simulate the sampling distribution of W in R and plot your results (adapt code from the previous exercise). Check that the simulated mean and the standard error are close to the theoretical mean and the standard error.
(c) Use your simulation to find $P(W < -1.5)$. Calculate an exact answer and compare.

```r
set.seed(13)
sims <- 10000
xbar <- numeric(sims)
ybar <- numeric(sims)
for(i in 1:sims){
  xbar[i] <- mean(rnorm(9, 7, 3))
  ybar[i] <- mean(rnorm(12, 10, 5))
}
W <- xbar - ybar
ggplot(data = data.frame(x = W), aes(x = x)) + geom_density() + theme_bw() + geom_vline(xintercept = mea
```




```r
mean(W)   # close to -3
```

```
## [1] -3.025
```

```r
sd(W)     # close to 1.76
```

```
## [1] 1.758
```

```r
mean(W < -1.5)
```

```
## [1] 0.8091
```

```r
# Exact answer
pnorm(-1.5, -3, sqrt(3^2/9 + 5^2/12))
```

```
## [1] 0.8035
```

Prob 14(a) answer: The exact sampling distribution of W is given by $W\tilde{\ }N(-3, 1.7559)$.

Prob 14(b) answer: See plot above.

Prob 14(c) answer: By our simulation, $P(W < -1.5) = 0.8091$. The exact value for $P(W < -1.5)$ is 0.8035. The simulated value has 0.6951% error.

Problem 15

Let $X_1, X_2, \ldots X_n$ be a random sample from N(0, 1). Let W $= X_1^2 + X_2^2 + \ldots + X_n^2$.

Describe the sampling distribution of W by running a simulation, using n = 2. What is the mean and variance of the sampling distribution of W? Repeat using n = 4, n = 5. What observations or conjectures do you have for general n?

```
set.seed(13)
sims <- 10000
WE2 <- numeric(sims)
WE4 <- numeric(sims)
WE5 <- numeric(sims)

for (i in 1:sims) {
  WE2[i] <- sum(rnorm(2)^2)
  WE4[i] <- sum(rnorm(4)^2)
  WE5[i] <- sum(rnorm(5)^2)
}

mean(WE2)
```

```
## [1] 2.007
```

```
mean(WE4)
```

```
## [1] 3.951
```

```
mean(WE5)
```

```
## [1] 5.025
```

```
var(WE2)
```

```
## [1] 4.051
```

```
var(WE4)
```

```
## [1] 7.957
```

```
var(WE5)
```

```
## [1] 9.773
```

> **Prob 15 answer:** For n=2, the mean of W is **2.007 and the variance of W is** 4.0514**. For n=4, the mean of W is** 3.9511 and the variance of W is **7.9574**. For n=5, the mean of W is **5.0247 and the variance of W is** 9.7725**. The mean of W is approximately** n**. The variance of W is approximately** 2n**.

Problem 17

Let $X_1, X_2, \ldots X_{20} \overset{iid}{\sim} \text{Exp}(2)$. Let X $= \sum_{i=1}^{20} X_i$.

(a) Simulate the sampling distribution of X in R.
(b) From your simulation, find E [X] and Var[X].
(c) From your simulation, find P(X $\leq$ 10).

```
set.seed(13)
sims <- 10000
WE <- numeric(sims)
for(i in 1:sims){
  WE[i] <- sum(rexp(20, 2))
}
mean(WE)
```

```
## [1] 9.965
```

```
var(WE)
```

```
## [1] 4.901
```

```
mean(WE <= 10)
```

```
## [1] 0.5369
```

Prob 17b answer: The expected value for X is **9.9647** and the variance of X is **4.901**.

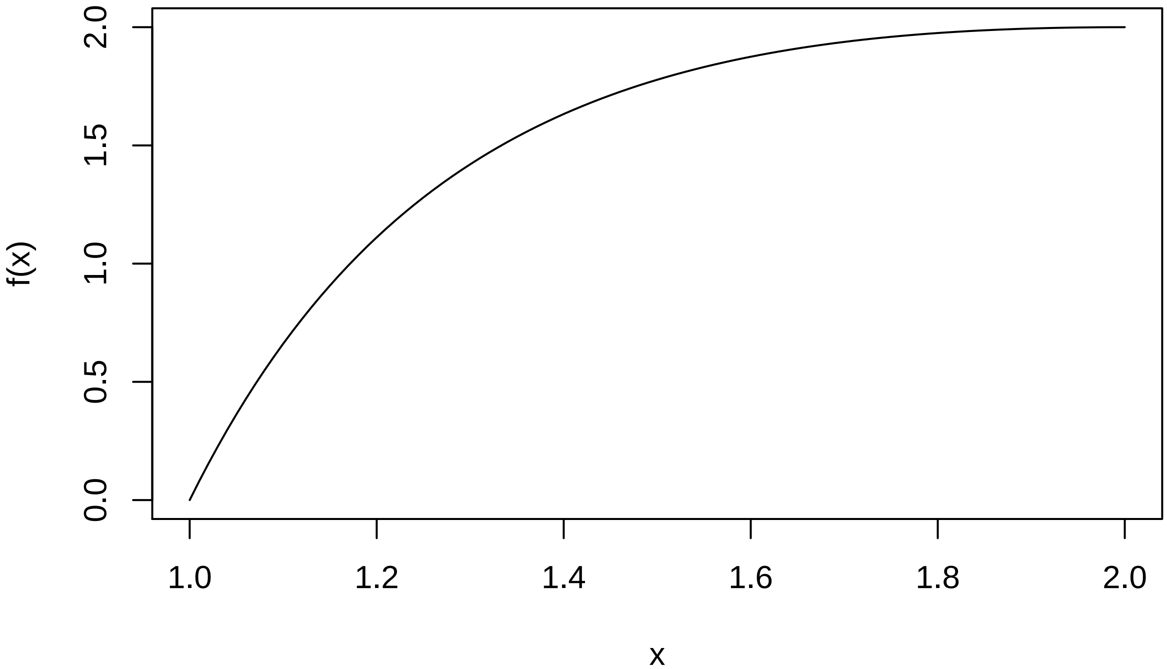
Prob 17c answer: The probability that X <= 10 is **0.5369**.

Problem 21

Let $X_1, X_2 \overset{iid}{\sim}$ F with corresponding pdf $f(x) = \frac{2}{x^2}, 1 \leq x \leq 2$.

(a) Find the pdf of X_{max}.
(b) Find the expected value of X_{max}.

```
f <- function(x){x * 2*(2 -2/x)^(2-1) * 2/x^2}
curve(f, from=1,to=2)
```




```
ans <- integrate(f, 1, 2)$val
```

Prob 21a answer: The pdf of X_{max} is $f_{\mathbf{max}}(x) = n\,(2 - 2/x)^{n-1}\,2/x^2, 1 \le x \le 2$.

Prob 21b answer: The expected value of X_{max} is **1.5452**.