

Homework 04 - STAT440

Joseph Sepich (jps6444)

09/19/2020

```
set.seed(42)
```

Problem 1

Use R to simulate samples from a normal distribution. Let Y be a random variable with chi-squared distribution with 5 degrees of freedom.

Part a

Since we can only sample from the standard normal distribution, I would use the transformation that the Chi-square distribution is really the sum of squares of standard normal variables. The degrees of freedom is the number of standard normal random variables that make up the Chi-square distribution. In this case we would use the following transformation:

$$\chi^2 = \sum_{i=1}^5 Z_i^2$$

Part b

Variance of Y :

$$Var(Y) = \left(\int x f(x) dx \right)^2 - \left(\int x^2 f(x) dx \right)$$

Fourth moment of Y :

$$\int x^4 f(x) dx$$

We can approximate all these integrals using Monte Carlo, where $h(x) = x^n$ (depending on the moment) and $f(x)$ is the density of Y that we are sampling from. Therefore we can use our samples X_i from the standard normal and transform them (5 for 5 degrees of freedom) into a chi-square sample Y_i . These samples can give us the value $\frac{1}{n} \sum_{i=1}^n h(Y_i) = \frac{1}{n} \sum_{i=1}^n Y_i^k$, which can approximate the k^{th} moment integral.

Part c

Use R to estimate the above quantities using Monte Carlo with $N = 10,000$ samples and report the results.

```

n <- 10000
df <- 5
norm_samples <- matrix(rnorm(n*df), nrow=n, ncol=df)
norm_samples <- apply(norm_samples, c(1,2), function(x) {x^2})
chi_samples <- rowSums(norm_samples)

# variance
second <- sum(chi_samples^2) / n
expect_square <- (sum(chi_samples) / n) ^ 2
variance <- second - expect_square
print(variance)

```

```
## [1] 9.983276
```

```

# compare approx with sd function
print(sd(chi_samples)^2)

```

```
## [1] 9.984275
```

```

# fourth moment
fourth <- sum(chi_samples^4) / n
print(fourth)

```

```
## [1] 3449.642
```

Problem 2

Part a

Express the probability as an expectation and as an integral. For this we can use an indication function $g(x)$, which is 1 when $|\bar{X}_n - \mu| < \epsilon$ and 0 when it is not. ($|\bar{X}_n - \mu| \geq \epsilon$) We can use this indication function to express the probability as an expectation:

$$E[g(x)] = P(|\bar{X}_n - \mu| < \epsilon)$$

We represent this as an integral (formula for finding the first moment) where $|\bar{X}_n - \mu|$ is the value we plug in for x and $f(x)$ is the pdf of $|\bar{X}_n - \mu|$.

$$E[g(x)] = \int g(x)dP(x) = \int g(x)f(x)dx$$

Logically this makes sense as $P(|\bar{X}_n - \mu| < \epsilon)$ looks like the representation of a CDF, which is the integral we got.

Part b

Here $I_n = \int h(x)f(x)dx$ where $h(x)$ is the indicator function that we described above. This integral is in the form of an expectation of the indicator function, therefore all we need to do is sample from this indicator function random variable and find that sample mean, which is the Monte Carlo integration process.

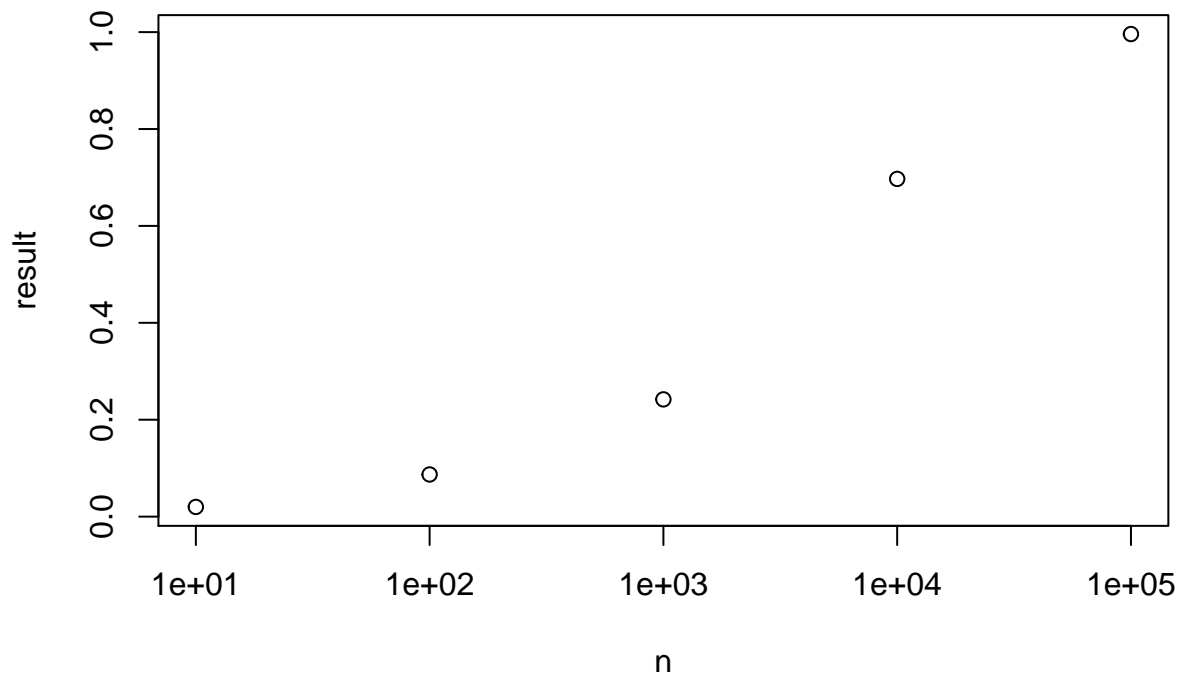
Part c

```
epsilon <- 0.01
m <- 1000
n <- c(10, 100, 1000, 10000, 100000)
result <- vector("numeric", length(n))

mu <- 2
sigma <- 1

for (i in 1:length(n)) {
  samples <- matrix(rnorm(n[i] * m, mu, sigma), nrow = m, ncol = n[i])
  samples <- rowSums(samples)
  x_n <- samples / n[i]
  comp_val <- abs(x_n - mu)
  result[i] <- sum(as.integer(comp_val < epsilon)) / m
}

plot(n, result, log='x')
```



Part d

As you can see in the plot above the correct behavior is shown. As the sample size of the sample mean gets larger, the expected value of our indicator function gets closer to 1, which means the numerical approximation

of the $P(|\bar{X}_n - \mu| < \epsilon)$ gets closer to 1. This is exactly what the limit states in the weak law of large numbers.

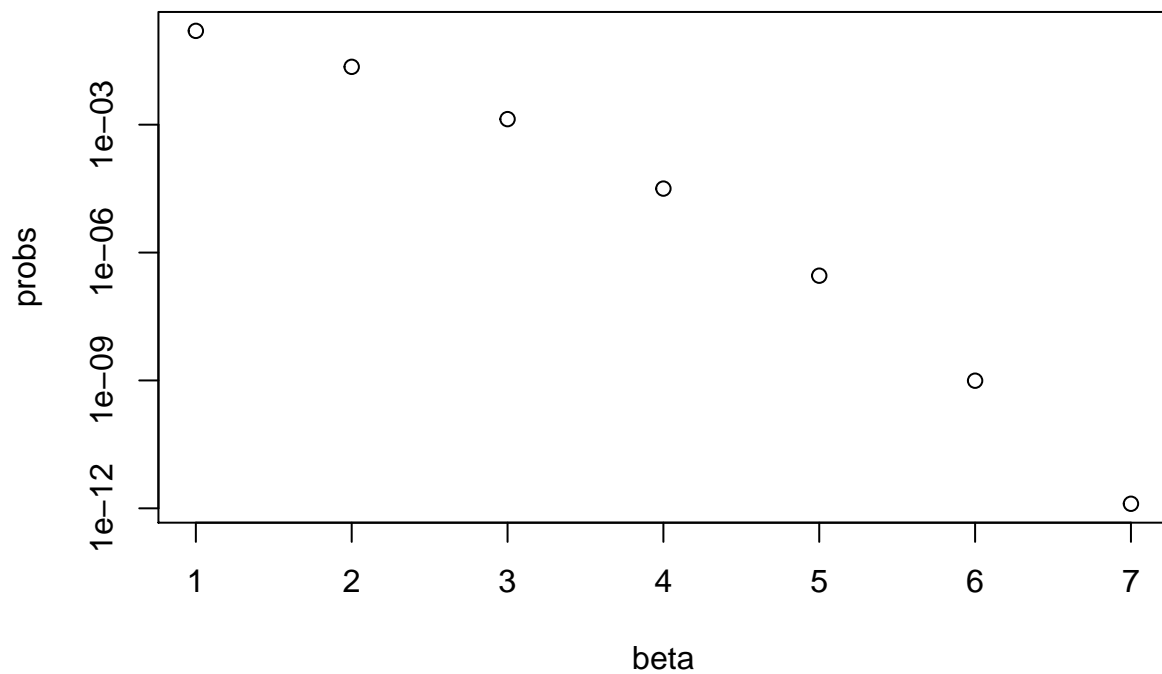
Problem 3

Part a

```
beta <- c(1:7)
mu <- 0
sigma <- 1

probs <- 1 - pnorm(beta, mu, sigma)

plot(beta, probs, log="y")
```



It would not be very reasonable to sample from the normal distribution and only take these large samples. These samples occur so sparsely that you would have to sample an incredible number of times. For example, when beta is 1, you could only use 15.8% of the samples to try to estimate the value. For beta of 2, it would be 2.3% and with beta of 7 you would be lucky to find 1 in 1,000,000 to use to estimate the value.

Part b

To find the pdf of $P(Z|Z > \beta)$ we will use the density version of Bayes theorem described where $S = Z > \beta$:

$$f_{Z|Z>\beta}(x) = \frac{f_Z(x)P(Z > \beta|Z = x)}{1 - \Phi(\beta)}$$

In the numerator we have $P(Z > \beta|Z = x)$, which is either 1 when $x > \beta$ or 0 otherwise, which makes sense, because you can't condition on the event that the value is greater than β unless it is. This gives us the density:

$$f_{Z|Z>\beta}(x) = \frac{f_Z(x)}{1 - \Phi(\beta)}$$

With the support $\beta < x < \infty$.

We can use this to find expectation:

$$\begin{aligned}\theta_\beta &= \int_\beta^\infty x f_{Z|Z>\beta}(x) dx = \int_\beta^\infty \frac{x f_Z(x)}{1 - \Phi(\beta)} dx = \frac{1}{1 - \Phi(\beta)} \int_\beta^\infty x f_Z(x) dx \\ \theta_\beta &= \frac{1}{1 - \Phi(\beta)} \int_\beta^\infty \frac{x}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{-1}{\sqrt{2\pi}(1 - \Phi(\beta))} (e^{-x^2/2})|_\beta^\infty \\ \theta_\beta &= \frac{e^{-\beta^2/2}}{\sqrt{2\pi}(1 - \Phi(\beta))}\end{aligned}$$

Part c

We can use a rejection sampling approach to sample values from the distribution to estimate the expectation using the sample mean. The candidate density to be used will be that of the standard normal. Recall that we want to find a value that fits the following:

$$\frac{f(x)}{g(x)} \leq c$$

To do this we can define $h(x) = \frac{f(x)}{g(x)}$,

$$h(x) = \frac{f(x)}{g(x)} = \frac{\frac{f_Z(x)}{1 - \Phi(\beta)}}{f_Z(x)} = \frac{1}{1 - \Phi(\beta)}$$

This means our scaling value $c = \frac{1}{1 - \Phi(\beta)}$. For a value of $\beta = 6$ we find our probability of acceptance to be:

```
beta <- 6
c <- 1 / (1 - pnorm(6))
1 / c
```

```
## [1] 9.865877e-10
```

This makes the standard normal a very poor density to use since we would have to wait for about 1013594635 samples to get 1 accepted. This is so poor, since the support of the conditional density only exists in the area greater than β while the standard normal support covers negative infinity to infinity. Since the values in the desired support will be drawn so rarely from the candidate function this makes it a very poor choice.