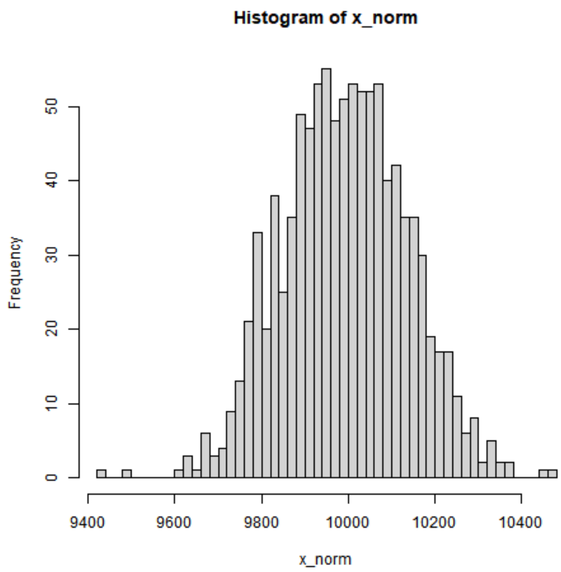
Homework 2: Gaussian Annulus and Random Projection

MAS480 Advanced Mathematics for Data Science

20200130 Yujun Kim

Let $d = 10^4, n = 10^3$.

1-(c) Draw a histogram of $||X_i||^2, i = 1, \cdots, n$ with at least 50 bins.



1-(d) What is the sample mean and sample standard deviation of $||X_i||^2$? Compare them with the exact expected values

```
> # 1-(d) Sample mean and Sample standard deviation
> mean(x_norm)
[1] 10001.02
> sd(x_norm)
[1] 139.5385
```

`x_mean` in above code is vector with $||X_i||^2$ as each element. The mean is $\mu = 10001.02$ and the standard deviation is $\sigma = 139.54$. Suppose $Z \sim N(0, 1)$. Then the moment generating function is

$$M_z(t) = e^{\frac{1}{2}t^2}$$

$$E[Z^2] = M_z''(0) = 1$$

$$V[Z^2] = E[Z^4] - E[Z^2]^2 = M_z^{(4)}(0) - 1^2 = 3 - 1 = 2$$

Thus, $||X_i||^2$ as sum of n IID standard normal, the expected mean is 10000 and the expected standard deviation is $\sqrt{20000} = 141.42$. Hence, the experimental result is close enough to the expected value.

2-(d) Draw a histogram of $r_i, i = 1, \cdots, n$ with at least 50 bins.

Figure1 is the histogram of r_i with m=100, 300, 500 respectively. Note that x scale varies.

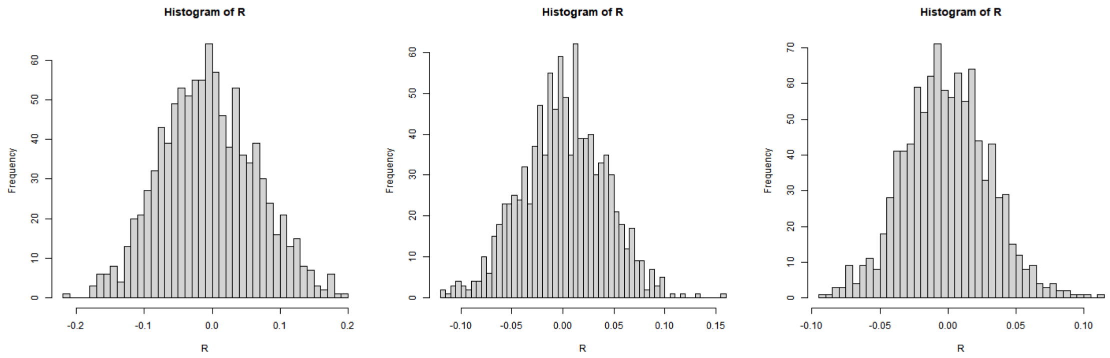


Figure 1 r_i with m=100, 300, 500 Respectively

2-(e) What is the ratio of r_i with $|r_i| \leq 0.1$?

See the result of execution below. The ratio is 0.841, 0.986, 0.998 for m=100, 300, 500 respectively.

```
> # 2-(e) m=100
> sum(abs(R) <= 0.1)/n
[1] 0.841

> # 2-(e) m=300
> sum(abs(R) <= 0.1)/n
[1] 0.986

> # 2-(e)
> sum(abs(R) <= 0.1)/n
[1] 0.998
```

This matches with theorem 9 of the lecture slide as the probability of random X having absolute value of r greater than 0.1 decrease exponentially.