Groupwork 9-5-2017

student

9/5/2017

Part 1:

X_1 is Normal(2,4), X_2 is Normal(5,1), and X_3 is Normal(0,1). X_1, X_2 and X_3 are independent.

Let $Y = aX_1 + bX_2 + cX_3$.

1. What is the expected value of Y?

$E[Y] = 2a + 5b$

2. What is the variance of Y?

$Var[Y] = 4a^2 + b^2 + c^2$

3. Does Y follow a Normal distribution?

Yes

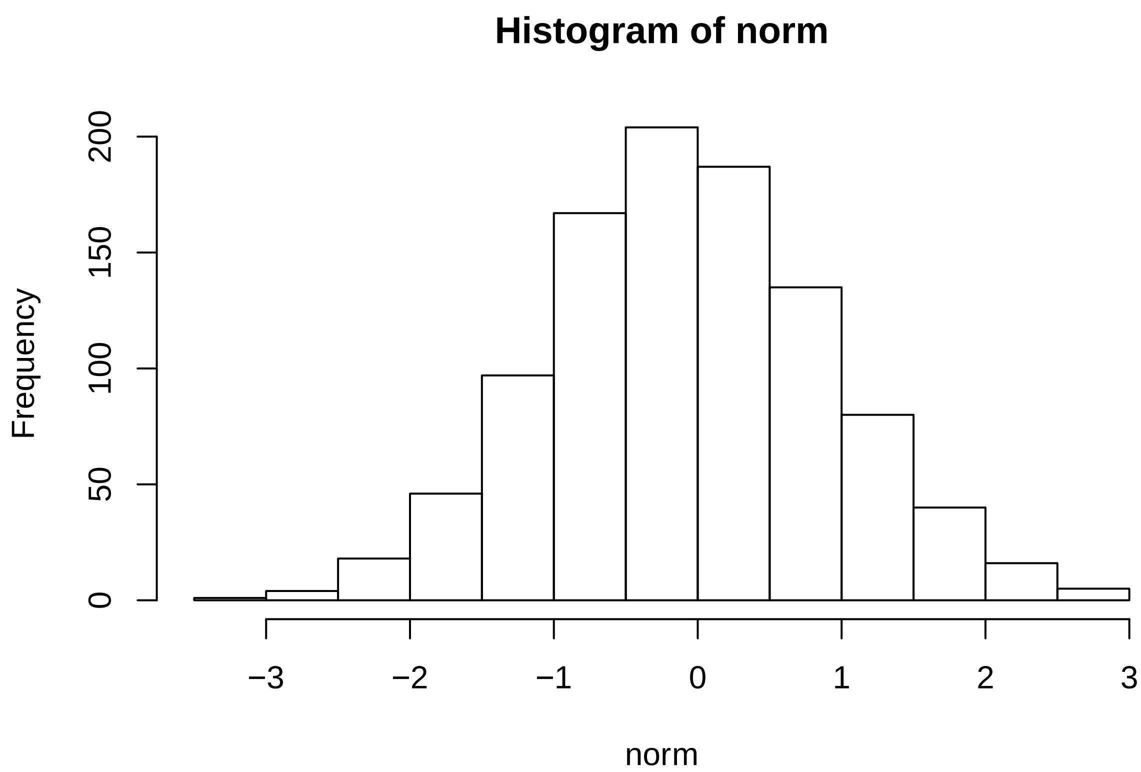
4. Does Y^2 follow a Normal distribution?

No

Part 2:

1. Generate 1000 samples from a Normal(0,1) distribution and save to a variable called `x1`. Plot a histogram of the samples using the `hist()` function. Describe the distribution.

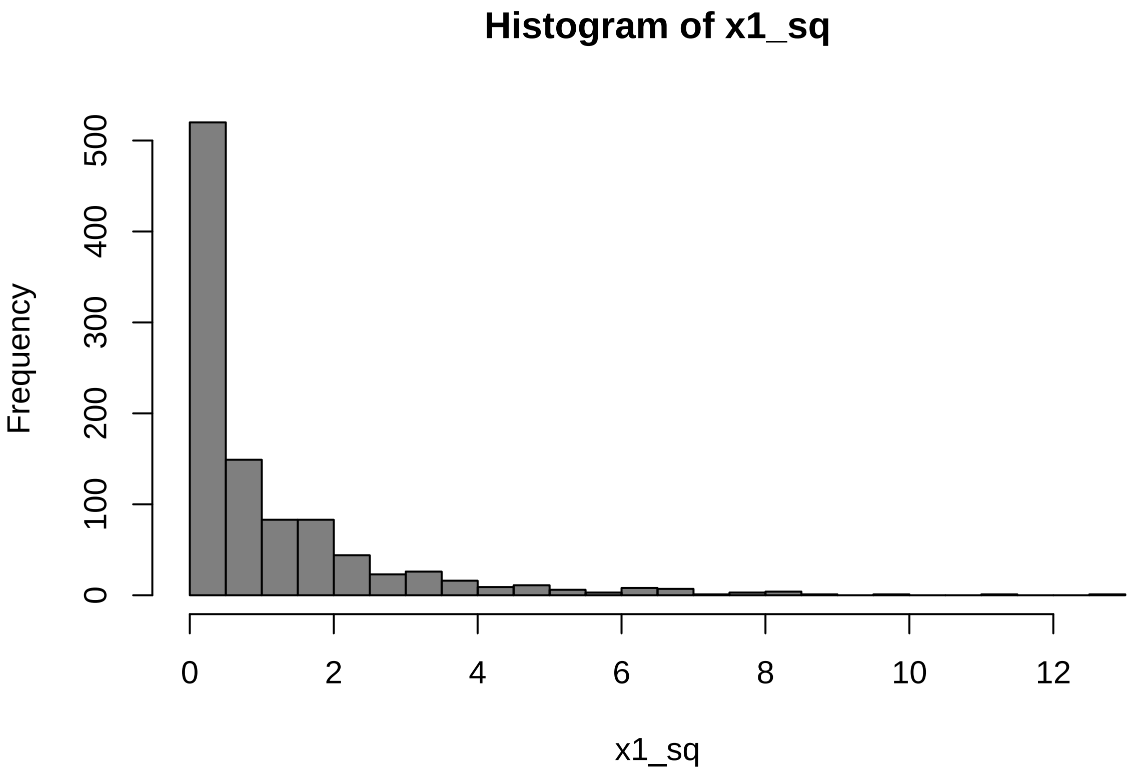
```r
set.seed(124)
norm <- rnorm(1000, mean = 0, sd = 1)
hist(norm)
```



It's normal

2. Square the samples and save to a variable called `x1_sq`. Plot a histogram using the `hist()` function with `breaks=30` and `col=adjustcolor("black", alpha.f=0.5)` (this will make the histogram bars transparent). Describe the distribution.

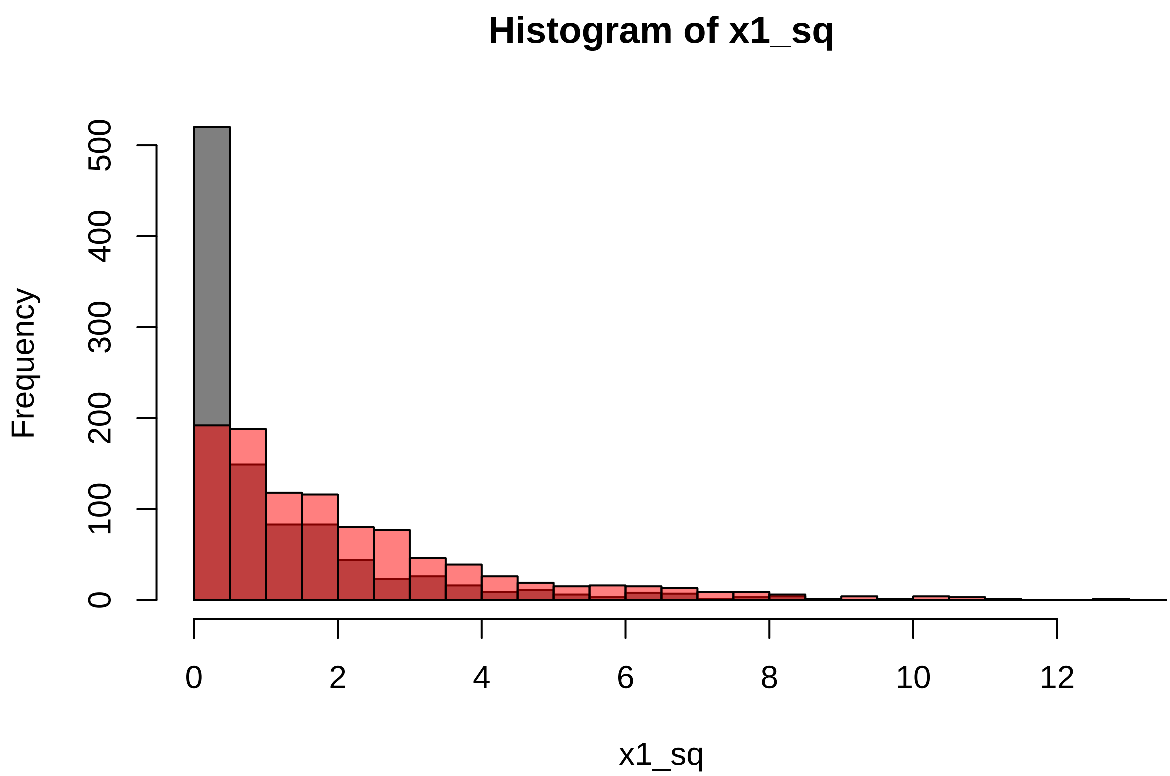
```
x1_sq <-  rnorm(1000, mean = 0, sd = 1) ** 2
hist(x1_sq, breaks=30, col=adjustcolor("black", alpha.f=0.5))
```



Concave distribution. Looks like a zipf distribution.

3. Now generate another set of 1000 samples from a Normal(0,1) and save to a variable called `x2`. Square these, add them to `x1_sq`, and save the result as `x1_x2_sq`. Plot a histogram with `breaks=30`, `col=adjustcolor("red", alpha.f=0.5)` and `add=TRUE` (this will overlay the histogram over the existing one). (Because this is a separate chunk, you'll need to copy and paste the code used to create your original histogram and include it here.) Describe the distribution.

```
x2 <- rnorm(1000, mean = 0, sd = 1)
x1_x2_sq <- (x2 ** 2) + x1_sq
hist(x1_sq, breaks=30, col=adjustcolor("black", alpha.f=0.5))
hist(x1_x2_sq, breaks=30, col=adjustcolor("red", alpha.f=0.5), add=TRUE)
```



Much more even than `x1_sq`

4. Finally, generate one more set of 1000 samples from a Normal(0,1) and save to a variable called `x3`. Square these, add them to `x1_x2_sq`, and save the result as `x1_x2_x3_sq`. Plot a histogram with `breaks=30`, `col=adjustcolor("blue", alpha.f=0.5)` and `add=TRUE`.

[Insert chunk here]

5. Compare and contrast the distributions from parts 1-3. How do their means compare? Their variances? You can check your intuition by computing the mean and variance of `x1_sq`, `x1_x2_sq` and `x1_x2_x3_sq`.

[Insert text here]

6. In general, what happens to the variance of a random variable when we add another independent random variable to it? Check that your answer agrees with your response to part 1, question 2 and part 2, question 4.

[Insert text here]

Part 3:

1. Generate 1000 samples from a Chi-squared distribution with 1 degree of freedom sand save to a variable called `y_1df`. Plot overlaid histograms of `x1_sq` and `y_1df`, using transparency and two different colors as above. Use the same number of breaks for both histograms. Compare the two distributions.

[Insert chunk here]

[Insert text here]

2. Generate 1000 samples from a Chi-squared distribution with 2 degrees of freedom sand save to a variable called `y_2df`. Plot overlaid histograms of `x1_x2_sq` and `y_2df`, using transparency and two different colors as above. Use the same number of breaks for both histograms. Compare the two distributions.

[Insert chunk here]

[Insert text here]