2.2 CHALLENGES

Sunday, August 14, 2022 3:32 PM

Myroslava Sánchez Andrade A01730712 | 14/08/2022

2.2.2 CHALLENGE

CALCULATE THE MEAN AND STANDARD DEVIATION OF THIS SIMULATED UNIFORM x VARIABLE. HOW CLOSE WERE THESE MEAN AND STANDARD DEVIATION FROM THE THEORETICAL MEAN AND STANDARD DEVIATION?

```
# Calculating the theoretical mean and standard deviation
x_mean_theoretical = (0 + 40)/2
x_std_theoretical = math.sqrt((40 - 0)**2/12)
print('Theoretical')
print(f'Mean: {x_mean_theoretical} Std: {x_std_theoretical}')
print('Simulated')
print(f'Mean: {x.mean()} Std: {x.std()}')
```

```
Theoretical
Mean: 20.0 Std: 11.547005383792516
Simulated
Mean: 20.092600526638293 Std: 11.479747505148312
```

A. Calculate the mean and standar deviation of this simulated uniform x variable. How close where these mean and standard deviation from the theoreticla mean and standard deviation?

```
# Calculating the theoretical mean and standard deviation
x_mean_theoretical = (0 + 40)/2
x_std_theoretical = math.sqrt((40 - 0)**2/12)
print('Theoretical')
print(f'Mean: {x_mean_theoretical} Std: {x_std_theoretical}')
print('Simulated')
print(f'Mean: {x.mean()} Std: {x.std()}')
```

```
Theoretical
Mean: 20.0 Std: 11.547005383792516
Simulated
Mean: 20.092600526638293 Std: 11.479747505148312
```

IT IS NOT EXACTLY THE SAME RESULT, BUT BOTH, THE THEORETICAL AND SIMULATED VALUES ARE CLOSE TO EACH OTHER. THE DIFFERENCE OCCURS DUE TO THE RANDOM PROCESSES IN THIS SIMULATION.

B. Interpret the above histogram, what do you see? Explain with your words. THE HISTOGRAM REPRESENTS A UNIFORM DISTRIBUTION, WHICH MEANS THAT ALL OUTCOMES (NUMBERS FROM 0 - 40) ARE EQUALLY LIKELY. IT LOOKS APROXIMATELY LIKE A RECTANGLE, WHERE THE BASE IS 40 AND THE HEIGHT APPROXIMATELY 0.025

2.2.3 CHALLENGE

2.2.3 CHALLENGE

A. Calculate the mean and standard deviation of the xmean variable. How similar or different are these values from the mean and standar deviation of x?

```
print('X')
print(f'Mean: {x.mean()} Std: {x.std()}')
print('xmean')
print(f'Mean: {xmean.mean()} Std: {xmean.std()}')
```

```
X
Mean: 20.092600526638293 Std: 11.479747505148312
xmean
Mean: 20.02561999958429 Std: 2.291276405175547
```

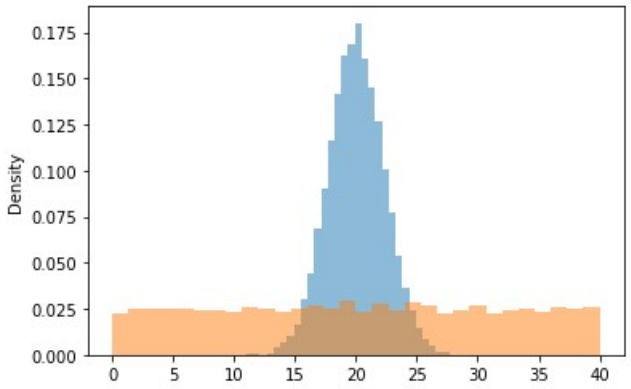
WE CAN CLEARLY REALIZE THAT THERE IS BASICALLY NO DIFFERENCE BETWEEN THE MEAN, THE MAIN DIFFERENCE IS BETWEEN THE STANDARD DEVIATIONS. THIS IS CAUSED BECAUSE OF THE AGGREGATION OF EACH GROUP BY ITS MEAN.

B. What do you see in the above histogram? Does the histogram of sample means of uniform random variable look like a uniform distributed variable? Do you see a logic? Briefly explain what you think.

NO, THE HISTOGRAM LOOKS LIKE A NORMAL DISTRIBUTION AND THIS HAPPENS BEACUSE WE ARE PLOTTING THE MEAN OF THE GROUPS, WHICH MAKES VALUES CLOSER TO THE OVERALL MEAN.

Now plot both variables in the same plot; the original x uniform variable and the xsample variable (the sample means of x).

```
plt.clf()
xmean.plot()
x.plot()
plt.show()
```



C. What do you see? How the histogram of x differs from the histogram of the sample of x? THE BLUE HISTOGRAM REPRESENTS A NORMAL DISTRIBUTION (BELL-SHAPED), MEANWHILE THE ORANGE HISTOGRAM REPRESENTS A UNIFORM DISTRIBUTION.

2.2.5 CHALLENGE

A. Calculate the mean and standard deviation of this random y variable. How similar or different are these values from the mean and standard deviation of the theorical mean (20) and theorical standard deviation (10)?

```
print('Y')
print(f'Mean: {y.mean()} Std: {y.std()}')
print("Theoretical")
print("Mean: 20, Std: 10")
```

```
Y
Mean: 20.004639442177098 Std: 10.01583027205781
Theoretical
Mean: 20, Std: 10
```

BOTH RESULTS ARE QUITE SIMILAR BECAUSE WE ARE GENERATING THE NORMAL DISTRIBUTION FROM THE SAME VALUES.

B. What do you see? Interpet the histogram with your words.

GIVEN THIS NORMAL DISTRIBUTION, WE CAN OBSERVE THAT THE VALUES ARE CLOSE TO THE PROVIDED MEAN.

2.2.6 CHALLENGE

A. Calculate the mean and the standard deviation of the ymean variable. How similar or different are these values from the mean and standard deviation of y?

```
print('Y')
print(f'Mean: {y.mean()} Std: {y.std()}')
print('ymean')
print(f'Mean: {ymean.mean()} Std: {ymean.std()}')
```

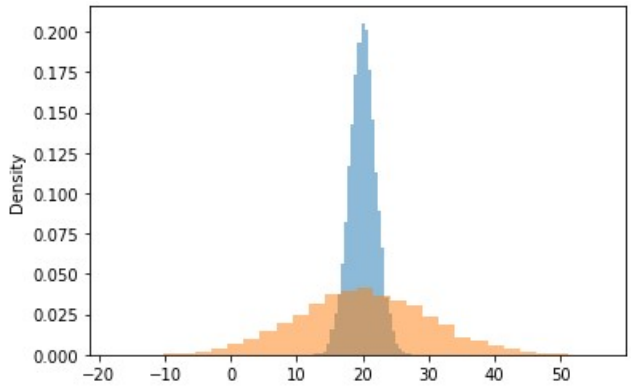
```
Y
Mean: 20.004639442177098 Std: 10.01583027205781
ymean
Mean: 20.00386712188444 Std: 1.9834702948105558
```

WE CAN CLEARLY REALIZE THAT THERE IS BASICALLY NO DIFFERENCE BETWEEN THE MEAN, THE MAIN DIFFERENCE IS BETWEEN THE STANDARD DEVIATIONS. THIS IS CAUSED BECAUSE OF THE AGGREGATION OF EACH GROUP BY ITS MEAN.

B. Interpret the above histogram of the samples of y.

THE HISTOGRAM REPRESENTS A NORMAL DISTRIBUTION, WHERE THE STANDARD DEVIATION IS SMALLER THAN THE ORIGINAL ONE AND THIS IS CAUSED BECAUSE OF THE GROUPING, WHICH MAKES VALUES TO BE CLOSER TO THE MEAN.

```
plt.clf()
ymean.plot()
y.plot()
plt.show()
```



C. What do you see? Compare both histograms, briefly explain what you think that happened.

THE GROUPING IN THE LAST NORMAL DISTRIBUTION CAUSES IT TO HAVE A HIGHER DENSITY AROUND THE MEAN AND LESS SPREAD ACROSS THE X-VALUES.