|  |  |
| --- | --- |
| Activity | Data Type |
| Number of beatings from Wife | Integer |
| Results of rolling a dice | integer |
| Weight of a person | Float or Decimal |
| Weight of Gold | Float or Decimal |
| Distance between two places | Float or Decimal |
| Length of a leaf | Float or Decimal |
| Dog's weight | Float or Decimal |
| Blue Color | String |
| Number of kids | integer |
| Number of tickets in Indian railways | integer |
| Number of times married | integer |
| Gender (Male or Female) | string |

Q1) Identify the Data type for the Following:

Q2) Identify the Data types, which were among the following

Nominal, Ordinal, Interval, Ratio.

|  |  |
| --- | --- |
| Data | Data Type |
| Gender | Nominal |
| High School Class Ranking | ordinal |
| Celsius Temperature | interval |
| Weight | Ratio |
| Hair Color | Nominal |
| Socioeconomic Status | Ordinal |
| Fahrenheit Temperature | Interval |
| Height | Ratio |
| Type of living accommodation | Nominal |
| Level of Agreement | Ordinal |
| IQ(Intelligence Scale) | Ratio |
| Sales Figures | Ratio |
| Blood Group | Nominal |
| Time Of Day | Interval |
| Time on a Clock with Hands | Interval |
| Number of Children | Ratio |
| Religious Preference | Nominal |
| Barometer Pressure | Interval |
| SAT Scores | Ratio |
| Years of Education | Ratio |

Q3) Three Coins are tossed, find the probability that two heads and one tail are obtained?

Ans = Now, add the probabilities of these three combinations to get the total probability of getting two heads and one tail:

Total Probability = (1/8) + (1/8) + (1/8) = 3/8

So, the probability of getting two heads and one tail when three coins are tossed is 3/8.

1.H,H,T

2.H,T,H

3.T,H,H

Q4) Two Dice are rolled, find the probability that sum is

1. Equal to 1
2. Less than or equal to 4
3. Sum is divisible by 2 and 3

Ans = Probability = [(5/36) + (1/36)] = 6/36 = 1/6

So, the probabilities are: a) Probability that the sum is equal to 1: 1/36 b) Probability that the sum is less than or equal to 4: 1/6 c) Probability that the sum is divisible by both 2 and 3: 1/6

Q5) A bag contains 2 red, 3 green and 2 blue balls. Two balls are drawn at random. What is the probability that none of the balls drawn is blue?

Ans = This is equivalent to selecting 2 balls out of 7 without regard to order, which is a combination.

Total number of ways to choose 2 balls out of 7 (regardless of color) = C(7, 2) = 21 ways.

There are 5 non-blue balls (2 red and 3 green) in the bag.

Number of ways to choose 2 non-blue balls out of 5 = C(5, 2) = 10 ways.

Now,the probability of drawing 2 non-blue ball:

Probability = (Number of ways to draw 2 non-blue balls) / (Total number of ways to draw 2 balls)

Probability = 10 / 21

So, the probability that none of the balls drawn is blue is 10/21.

Q6) Calculate the Expected number of candies for a randomly selected child

Below are the probabilities of count of candies for children (ignoring the nature of the child-Generalized view)

|  |  |  |
| --- | --- | --- |
| CHILD | Candies count | Probability |
| A | 1 | 0.015 |
| B | 4 | 0.20 |
| C | 3 | 0.65 |
| D | 5 | 0.005 |
| E | 6 | 0.01 |
| F | 2 | 0.120 |

Child A – probability of having 1 candy = 0.015.

Child B – probability of having 4 candies = 0.20

Ans = Expected Number of Candies = (Probability of Child A having 1 candy \* 1) + (Probability of Child B having 4 candies \* 4) + (Probability of Child C having 3 candies \* 3) + (Probability of Child D having 5 candies \* 5) + (Probability of Child E having 6 candies \* 6) + (Probability of Child F having 2 candies \* 2)

Expected Number of Candies = (0.015 \* 1) + (0.20 \* 4) + (0.65 \* 3) + (0.005 \* 5) + (0.01 \* 6) + (0.120 \* 2)

Expected Number of Candies = 0.015 + 0.80 + 1.95 + 0.025 + 0.06 + 0.24

Expected Number of Candies = 4.095

So, the expected number of candies for a randomly selected child is 4.095 candies.

Q7) Calculate Mean, Median, Mode, Variance, Standard Deviation, Range & comment about the values / draw inferences, for the given dataset

* For Points,Score,Weigh>

Find Mean, Median, Mode, Variance, Standard Deviation, and Range and also Comment about the values/ Draw some inferences.

Ans = let's calculate the requested statistics:

1. Mean (Average):
   * Mean Points = Sum of Points / Number of Data Points
   * Mean Score = Sum of Score / Number of Data Points
   * Mean Weigh = Sum of Weigh / Number of Data Points

Mean Points = (115.09 / 32) ≈ 3.5953 Mean Score = (98.675 / 32) ≈ 3.0836 Mean Weigh = (571.16 / 32) ≈ 17.8488

1. Median (Middle Value): To find the median, we need to first order the data.

Median Points = 3.44 (middle value) Median Score = 3.695 (middle value) Median Weigh = 17.71 (middle value)

1. Mode (Most Frequent Value): The mode is the most frequent value(s) in the dataset.
   * There is no mode for Points because all values are unique.
   * There is no mode for Score because all values are unique.
   * Mode Weigh = 17.02 (appears twice)
2. Variance: Variance measures the spread of the data.
   * Variance Points = Sum of [(Data Point - Mean Points)^2] / (Number of Data Points - 1)
   * Variance Score = Sum of [(Data Point - Mean Score)^2] / (Number of Data Points - 1)
   * Variance Weigh = Sum of [(Data Point - Mean Weigh)^2] / (Number of Data Points - 1)

Variance Points ≈ 0.289 Variance Score ≈ 1.141 Variance Weigh ≈ 3.193

1. Standard Deviation: Standard deviation is the square root of variance.
   * Standard Deviation Points ≈ 0.537
   * Standard Deviation Score ≈ 1.068
   * Standard Deviation Weigh ≈ 1.786
2. Range (Difference between the Maximum and Minimum Values):
   * Range Points = 4.93 - 2.76 = 2.17
   * Range Score = 5.424 - 1.513 = 3.911
   * Range Weigh = 22.9 - 14.5 = 8.4

Comments and Inferences:

* The mean Points is approximately 3.5953, mean Score is approximately 3.0836, and mean Weigh is approximately 17.8488.
* The medians for Points, Score, and Weigh are 3.44, 3.695, and 17.71, respectively.
* There is no mode for Points and Score as all values are unique. The mode for Weigh is 17.02, which appears twice.
* The variance and standard deviation for Weigh are higher than those for Points and Score, indicating greater variability in Weigh.
* The range for Weigh (8.4) is larger than the ranges for Points (2.17) and Score (3.911), suggesting greater variability in Weigh values.

**Use Q7.csv file**

Q8) Calculate Expected Value for the problem below

1. The weights (X) of patients at a clinic (in pounds), are

108, 110, 123, 134, 135, 145, 167, 187, 199

Assume one of the patients is chosen at random. What is the Expected Value of the Weight of that patient?

Ans = Weights (X) = [108, 110, 123, 134, 135, 145, 167, 187, 199]

The formula for calculating the expected value (mean) is:

Expected Value (μ) = (Sum of all values) / (Number of values)

Expected Value (μ) = (108 + 110 + 123 + 134 + 135 + 145 + 167 + 187 + 199) / 9

Expected Value (μ) = 1293 / 9

Expected Value (μ) ≈ 143.67 pounds

So, the expected value of the weight of a randomly chosen patient is approximately 143.67 pounds.

**Q9) Calculate Skewness, Kurtosis & draw inferences on the following data**

**Cars speed and distance**

**Use Q9\_a.csv**

**SP and Weight(WT)**

**Use Q9\_b.csv**

**Ans = Speed (SP):** The formula for skewness is:

Skewness = (3 \* (Mean - Median)) / Standard Deviation

The formula for kurtosis is:

Kurtosis = (Sum of ((X - Mean) ^ 4) / (N \* (Standard Deviation ^ 4))) - 3

Where:

* X is each data point.
* Mean is the mean (average) of the data.
* Median is the median of the data.
* Standard Deviation is the standard deviation of the data.
* N is the number of data points.

First, let's calculate the skewness and kurtosis for "Speed."

**Speed (SP):**

Mean (μ\_SP) = (Sum of all Speed values) / (Number of Speed values)

Standard Deviation (σ\_SP) = Square root of Variance

Variance (σ^2\_SP) = (Sum of ((Speed - Mean) ^ 2)) / (Number of Speed values)

Median (Mid-Range) = (Minimum Speed + Maximum Speed) / 2

Let's calculate these values:

Mean (μ\_SP) ≈ 15.8 Standard Deviation (σ\_SP) ≈ 4.14 Variance (σ^2\_SP) ≈ 17.16 Median ≈ 15.5

Now, let's calculate skewness for Speed:

Skewness\_SP = (3 \* (μ\_SP - Median)) / σ\_SP Skewness\_SP = (3 \* (15.8 - 15.5)) / 4.14 Skewness\_SP ≈ 0.2184

Next, let's calculate kurtosis for Speed:

Kurtosis\_SP = (Sum of ((SP - μ\_SP) ^ 4) / (N \* (σ\_SP ^ 4))) - 3 Kurtosis\_SP ≈ 1.856

**Distance (WT):** Now, let's calculate the skewness and kurtosis for "Distance."

Mean (μ\_WT) ≈ 42.98 Standard Deviation (σ\_WT) ≈ 24.32 Variance (σ^2\_WT) ≈ 590.28 Median ≈ 36

Now, let's calculate skewness for Distance:

Skewness\_WT = (3 \* (μ\_WT - Median)) / σ\_WT Skewness\_WT = (3 \* (42.98 - 36)) / 24.32 Skewness\_WT ≈ 0.835

Next, let's calculate kurtosis for Distance:

Kurtosis\_WT = (Sum of ((WT - μ\_WT) ^ 4) / (N \* (σ\_WT ^ 4))) - 3 Kurtosis\_WT ≈ 2.052

**Q10) Draw inferences about the following boxplot & histogram**





Ans = I can’t solve plz help it

**Q11)** Suppose we want to estimate the average weight of an adult male in Mexico. We draw a random sample of 2,000 men from a population of 3,000,000 men and weigh them. We find that the average person in our sample weighs 200 pounds, and the standard deviation of the sample is 30 pounds. Calculate 94%,98%,96% confidence interval?

Ans = Confidence Interval=Sample Mean±(Critical Value×Sample Size​Standard Deviation​)

Where:

* Sample Mean: 200 pounds
* Standard Deviation of the Sample: 30 pounds
* Sample Size: 2,000
* Critical Value depends on the desired confidence level and degrees of freedom (df).

Degrees of Freedom (df) = Sample Size - 1 = 2,000 - 1 = 1,999

Let's calculate the critical values for each confidence level and then calculate the confidence intervals:

1. 94% Confidence Interval:
   * Critical Value for 94% confidence level with df = 1,999 is approximately 1.645 (you can look this up in a t-table or use a calculator).

Now, calculate the margin of error:

Margin of Error = 1.645 \* (30 / √2,000) ≈ 3.663

94% Confidence Interval = 200 ± 3.663 = [196.337, 203.663]

1. 98% Confidence Interval:
   * Critical Value for 98% confidence level with df = 1,999 is approximately 2.329 (you can look this up in a t-table or use a calculator).

Now, calculate the margin of error:

Margin of Error = 2.329 \* (30 / √2,000) ≈ 4.894

98% Confidence Interval = 200 ± 4.894 = [195.106, 204.894]

1. 96% Confidence Interval:
   * Critical Value for 96% confidence level with df = 1,999 is approximately 2.054 (you can look this up in a t-table or use a calculator).

Now, calculate the margin of error:

Margin of Error = 2.054 \* (30 / √2,000) ≈ 4.319

96% Confidence Interval = 200 ± 4.319 = [195.681, 204.319]

So, the confidence intervals for the average weight of adult males in Mexico are as follows:

* 94% Confidence Interval: [196.337, 203.663]
* 98% Confidence Interval: [195.106, 204.894]
* 96% Confidence Interval: [195.681, 204.319]

**Q12)** Below are the scores obtained by a student in tests

**34,36,36,38,38,39,39,40,40,41,41,41,41,42,42,45,49,56**

1. Find mean, median, variance, standard deviation.
2. What can we say about the student marks?
3. Ans = Find mean, median, variance, and standard deviation:

Mean (Average): To find the mean, sum all the scores and divide by the number of scores.

Mean = (34 + 36 + 36 + 38 + 38 + 39 + 39 + 40 + 40 + 41 + 41 + 41 + 41 + 42 + 42 + 45 + 49 + 56) / 18 Mean = 678 / 18 Mean = 37.67 (rounded to two decimal places)

Median (Middle Value): To find the median, first, arrange the scores in ascending order and then find the middle value.

Arranged scores: 34, 36, 36, 38, 38, 39, 39, 40, 40, 41, 41, 41, 41, 42, 42, 45, 49, 56

Since there are 18 scores, the median is the average of the 9th and 10th scores:

Median = (40 + 41) / 2 Median = 40.5

Variance: Variance measures the spread or dispersion of the data.

Variance = [(Sum of (X - Mean)^2) / N]

Variance = [(Σ(X - Mean)^2) / N]

Variance = [(Σ(34-37.67)^2 + (36-37.67)^2 + ... + (56-37.67)^2) / 18]

Variance ≈ 53.39 (rounded to two decimal places)

Standard Deviation: The standard deviation is the square root of the variance.

Standard Deviation = √Variance

Standard Deviation ≈ √53.39 ≈ 7.30 (rounded to two decimal places)

1. What can we say about the student's marks?

Based on the calculated statistics:

* The mean score is 37.67, which is close to the center of the score distribution.
* The median score is 40.5, indicating that half of the scores are above this value and half are below.
* The standard deviation is approximately 7.30, which suggests that the scores have moderate variability around the mean.

Q13) What is the nature of skewness when mean, median of data are equal?

Ans = In terms of skewness:

1. If the mean and median are equal, the skewness of the data is close to zero. This means there is little or no skewness in the distribution.
2. A skewness of zero indicates that the data is symmetrically distributed, and there is an equal balance between values on the left and right sides of the mean.
3. In a symmetric distribution, there is no tendency for the data to be skewed to the left (negatively skewed) or to the right (positively skewed). The data is evenly distributed around the central value (the mean).

Q14) What is the nature of skewness when mean > median ?

Ans = when the mean is greater than the median, it suggests that the data is positively skewed, with a tail extending to the right, indicating the presence of larger values in the dataset

Q15) What is the nature of skewness when median > mean?

Ans = , when the median is greater than the mean, it suggests that the data is negatively skewed, with a tail extending to the left, indicating the presence of smaller values in the dataset.

Q16) What does positive kurtosis value indicates for a data ?

Ans = a positive kurtosis value indicates that the data has heavier tails, is more peaked, and has a higher probability of extreme values compared to a normal distribution. It is one of the measures used to assess the shape and characteristics of a probability distribution.

Q17) What does negative kurtosis value indicates for a data?

Ans = a negative kurtosis value indicates that the data has lighter tails, is flatter, and has a lower probability of extreme values compared to a normal distribution. It is one of the measures used to assess the shape and characteristics of a probability distribution.

Q18) Answer the below questions using the below boxplot visualization.



What can we say about the distribution of the data?

What is nature of skewness of the data?

What will be the IQR of the data (approximately)?

1. Ans = **Distribution of the Data:**
   * If the box in the boxplot is symmetrically centered within the whiskers, it suggests a roughly symmetric distribution.
   * If the box is skewed to the left (toward the lower values) with a longer left whisker, it indicates negative (left) skewness.
   * If the box is skewed to the right (toward the higher values) with a longer right whisker, it indicates positive (right) skewness.
   * Outliers, if present, may also provide insights into the distribution.
2. **Nature of Skewness:**
   * If the median (middle line inside the box) is closer to the lower quartile (bottom edge of the box), it suggests negative (left) skewness.
   * If the median is closer to the upper quartile (top edge of the box), it suggests positive (right) skewness.
3. **IQR (Interquartile Range):**
   * The IQR is the range between the upper quartile (Q3) and the lower quartile (Q1).
   * It represents the middle 50% of the data and gives a measure of the spread of the data.

Q19) Comment on the below Boxplot visualizations?



Draw an Inference from the distribution of data for Boxplot 1 with respect Boxplot 2.

Ans = I cant’t solve it .plz help this qus.

Q 20) Calculate probability from the given dataset for the below cases

Data \_set: Cars.csv

Calculate the probability of MPG of Cars for the below cases.

MPG <- Cars$MPG

* 1. P(MPG>38)
  2. P(MPG<40)
  3. P (20<MPG<50)

Ans = Mean (μ) = Σ(MPG) / N Standard Deviation (σ) = √[Σ(MPG - μ)^2 / (N - 1)]

Once you have the mean and standard deviation, you can convert the MPG values to z-scores and use standard normal distribution tables or a calculator to find the probabilities.

a. P(MPG > 38):

1. Calculate the z-score for 38: z = (38 - μ) / σ
2. Use a standard normal distribution table or calculator to find P(Z > z).

b. P(MPG < 40):

1. Calculate the z-score for 40: z = (40 - μ) / σ
2. Use a standard normal distribution table or calculator to find P(Z < z).

c. P(20 < MPG < 50):

1. Calculate the z-scores for 20 and 50: z1 = (20 - μ) / σ z2 = (50 - μ) / σ
2. Use a standard normal distribution table or calculator to find P(z1 < Z < z2)

Q 21) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

1. Ans = **Visual Inspection:** a. Create a histogram of the MPG values from your dataset. b. Plot a normal probability plot (Q-Q plot) of the MPG values. In a Q-Q plot, if the data points approximately follow a straight line, it suggests normality.
2. **Statistical Test (Shapiro-Wilk test):** a. Null Hypothesis (H0): The data follows a normal distribution. b. Alternative Hypothesis (Ha): The data does not follow a normal distribution. c. Set your significance level (alpha), e.g., α = 0.05. d. Perform the Shapiro-Wilk test on the MPG data. If the p-value is less than your chosen alpha level, reject the null hypothesis.

# Load the dataset (replace 'data.csv' with your actual file path)

data <- read.csv("Cars.csv")

# Visual Inspection

hist(data$MPG, main="Histogram of MPG", xlab="MPG", ylab="Frequency")

qqnorm(data$MPG)

qqline(data$MPG, col = 2)

# Statistical Test (Shapiro-Wilk)

shapiro.test(data$MPG)

1. Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

Dataset: wc-at.csv

Ans = # Load the dataset

data <- read.table(text = "Waist,AT

74.75,25.72

72.6,25.89

81.8,42.6

83.95,42.8

74.65,29.84

71.85,21.68

80.9,29.08

83.4,32.98

63.5,11.44

73.2,32.22

71.9,28.32

75,43.86

73.1,38.21

79,42.48

77,30.96

68.85,55.78

75.95,43.78

74.15,33.41

73.8,43.35

75.9,29.31

76.85,36.6

80.9,40.25

79.9,35.43

89.2,60.09

82,45.84

92,70.4

86.6,83.45

80.5,84.3

86,78.89

82.5,64.75

83.5,72.56

88.1,89.31

90.8,78.94

89.4,83.55

102,127

94.5,121

91,107

103,129

80,74.02

79,55.48

83.5,73.13

76,50.5

80.5,50.88

86.5,140

83,96.54

107.1,118

94.3,107

94.5,123

79.7,65.92

79.3,81.29

89.8,111

83.8,90.73

85.2,133

75.5,41.9

78.4,41.71

78.6,58.16

87.8,88.85

86.3,155

85.5,70.77

83.7,75.08

77.6,57.05

84.9,99.73

79.8,27.96

108.3,123

119.6,90.41

119.9,106

96.5,144

105.5,121

105,97.13

107,166

107,87.99

101,154

97,100

100,123

108,217

100,140

103,109

104,127

106,112

109,192

103.5,132

110,126

110,153

112,158

108.5,183

104,184

111,121

108.5,159

121,245

109,137

97.5,165

105.5,152

98,181

94.5,80.95

97,137

105,125

106,241

99,134

91,150

102.5,198

106,151

109.1,229

115,253

101,188

100.1,124

93.3,62.2

101.8,133

107.9,208

108.5,208", header = TRUE, sep = ',')

# Shapiro-Wilk Normality Test

shapiro.test(data$AT) # Test for AT

shapiro.test(data$Waist) # Test for Waist

# Create Histograms and Q-Q Plots

par(mfrow = c(2, 2))

hist(data$AT, main = "Histogram of AT", xlab = "AT")

qqnorm(data$AT)

qqline(data$AT, col = 2)

hist(data$Waist, main = "Histogram of Waist", xlab = "Waist")

qqnorm(data$Waist)

qqline(data$Waist, col = 2)

Q 22) Calculate the Z scores of 90% confidence interval,94% confidence interval, 60% confidence interval

1. Ans = For a 90% confidence interval:
   * The confidence level is 90%, which means there is a 5% chance in each tail (totaling 10% in both tails).
   * The central portion, where the data falls, is 100% - 10% = 90%.
   * To find the Z-score for the middle 90% (45% on each side), you can use a Z-table or a calculator.
   * The Z-score for a 90% confidence interval is approximately ±1.645.
2. For a 94% confidence interval:
   * The confidence level is 94%, which means there is a 3% chance in each tail (totaling 6% in both tails).
   * The central portion, where the data falls, is 100% - 6% = 94%.
   * To find the Z-score for the middle 94% (47% on each side), you can use a Z-table or a calculator.
   * The Z-score for a 94% confidence interval is approximately ±1.88.
3. For a 60% confidence interval:
   * The confidence level is 60%, which means there is a 20% chance in each tail (totaling 40% in both tails).
   * The central portion, where the data falls, is 100% - 40% = 60%.
   * To find the Z-score for the middle 60% (30% on each side), you can use a Z-table or a calculator.
   * The Z-score for a 60% confidence interval is approximately ±0.84.

Q 23) Calculate the t scores of 95% confidence interval, 96% confidence interval, 99% confidence interval for sample size of 25

Ans = The formula for calculating the t-score for a confidence interval is:

�=�ˉ−���*t*=*n*​*s*​*x*ˉ−*μ*​

Where:

* �ˉ*x*ˉ is the sample mean.
* �*μ* is the population mean (which you may not know, but you can use it as 0 for some calculations).
* �*s* is the sample standard deviation.
* �*n* is the sample size.

For a 95% confidence interval with 24 degrees of freedom (sample size - 1 = 25 - 1 = 24):

1. Use a t-distribution table or calculator to find the t-score for a 95% confidence interval with 24 degrees of freedom. The critical t-value is approximately ±2.064.

For a 96% confidence interval with 24 degrees of freedom:

1. Use a t-distribution table or calculator to find the t-score for a 96% confidence interval with 24 degrees of freedom. The critical t-value is approximately ±2.171.

For a 99% confidence interval with 24 degrees of freedom:

1. Use a t-distribution table or calculator to find the t-score for a 99% confidence interval with 24 degrees of freedom. The critical t-value is approximately ±2.797.

These t-scores represent the critical values for the specified confidence intervals when you have a sample size of 25 and want to calculate confidence intervals for the mean of a population.

Q 24**)** A Government company claims that an average light bulb lasts 270 days. A researcher randomly selects 18 bulbs for testing. The sampled bulbs last an average of 260 days, with a standard deviation of 90 days. If the CEO's claim were true, what is the probability that 18 randomly selected bulbs would have an average life of no more than 260 days

Hint:

rcode 🡪 pt(tscore,df)

df 🡪 degrees of freedom

Ans = the t-distribution and the given information.

First, you need to calculate the t-score using the formula:

�=�ˉ−���*t*=*n*​*s*​*x*ˉ−*μ*​

Where:

* �ˉ*x*ˉ is the sample mean (260 days).
* �*μ* is the claimed population mean (270 days).
* �*s* is the sample standard deviation (90 days).
* �*n* is the sample size (18 bulbs).

Now, calculate the t-score:

�=260−2709018*t*=18​90​260−270​

�=−109018*t*=18​90​−10​

Next, you need to find the degrees of freedom (��*df*) which is �−1*n*−1 in this case:

��=18−1=17*df*=18−1=17

Now, you can use the t-distribution to find the probability that the t-score is less than or equal to the calculated value. In R code, you can use the **pt** function:

RCopy code

# Calculate the probability tscore <- -10 / (90 / sqrt(18)) # Calculate t-score df <- 17 # Degrees of freedom probability <- pt(tscore, df) # Print the probability probability

This will give you the probability that 18 randomly selected bulbs would have an average life of no more than 260 days if the CEO's claim were true.