

LAB SESSION 3 – DISTRIBUTIONS

Analytics Primer

SUMMARIZING A DISTRIBUTION WITH OUTLIERS

Example

- A marketing firm collected data on annual household incomes for Outland, NC. They surveyed all 182 households in the small town of Outland. Now imagine that one of the marketing firm's vice presidents really liked the city after collecting data from them and moves to the city. Now the data set has 183 households. The vice president's household's annual income is \$384,000. Recalculate the mean, median, standard deviation, skewness, and kurtosis.

CONTINUOUS PROBABILITY DISTRIBUTIONS

Standard Normal Distribution

Example

- Assuming a normal distribution with mean = 0 and s.d. = 1, find the following probabilities:

1. $P(z > 2) = 0.0228$

2. $P(z \leq 1.12) = 0.8686$

3. $P(-1.33 \leq z \leq 1.33) = 0.9082 - 0.0918 = 0.8164$

4. $P(z \geq 3.02) = 0.0013$

5. $P(z \leq 6.87) \approx 1$

6. $P(z \leq -6.87) \approx 0$

YEARS OF
PROFESSIONAL
EXPERIENCE

More Examples

- Assume new employees at a company have previous years of professional experience that follow a normal distribution with the mean is 5 and the s.d. is 2.5.

Example

1. What is the probability that a new employee has more than 5 years of previous experience?

$$z = \frac{x - \mu}{\sigma} = \frac{5 - 5}{2.5} = 0 \rightarrow 0.5$$

Example

2. What is the probability that a new employee has less than 2 years of previous experience?

$$z = \frac{x - \mu}{\sigma} = \frac{2 - 5}{2.5} = -1.2 \rightarrow 0.1151$$

Example

3. What is the probability that a new employee has between 1 and 7.5 years of previous experience?

$$z = \frac{7.5 - 5}{2.5} = 1 \rightarrow 0.8413 \quad z = \frac{1 - 5}{2.5} = -1.6 \rightarrow 0.0548$$

$$0.8413 - 0.0548 = 0.7865$$

Example

4. What is the 90th percentile of employee experience?
What is the 10th percentile?

$$1.28 = \frac{x - 5}{2.5} \rightarrow x = 8.2$$

$$-1.28 = \frac{x - 5}{2.5} \rightarrow x = 1.8$$

MILES DRIVEN PER DAY

More Examples

- Americans drive an average of 41 miles per day with a standard deviation of 9. Assume that these miles follow a Normal distribution. Use this information to answer the next three questions

Example

1. What is the probability a random driver drives more than 60 miles in a day?

$$z = \frac{x - \mu}{\sigma} = \frac{60 - 41}{9} = 2.11 \rightarrow 0.0174$$

Example

2. What is the probability that a random driver drives less than 10 miles in a day?

$$z = \frac{x - \mu}{\sigma} = \frac{10 - 41}{9} = -3.44 \rightarrow 0.0003$$

Example

3. What is the interquartile range (IQR) for this data?

$$0.67 = \frac{x - 41}{9} \rightarrow x = 47.07 \quad -0.67 = \frac{x - 41}{9} \rightarrow x = 34.93$$

$$47.07 - 34.93 = 12.14$$