STA13: Elementary Statistics

Lecture 11
Book Section 4.5 - 4.6

Dmitriy Izyumin

February 7 2018

Standardization

There are infinitely many normal distributions.

Any normal variable X can be transformed into a standard normal variable Z.

If X is Normal(μ, σ), then $Z = \frac{X-\mu}{\sigma}$ is Normal(0, 1)

- start with X
- subtract μ
- divide by σ
- get Z

This lets us compute probabilities of any normal distribution using the standard normal distribution.

Typical Problem 1

Start with $X \sim N(\mu, \sigma)$.
Need to find a probability in terms of X.

1. Rewrite as a probability in terms of Z

2. Use the z-table to find the probability

Typical Problem 2

Start with $X \sim N(\mu, \sigma)$.
Need to find the k^{th} percentile of $N(\mu, \sigma)$

1. Compute the k^{th} percentile of $N(0, 1)$
2. Rescale it to get the percentile of $N(\mu, \sigma)$

Example 1

The height of female students entering a university has a normal distribution with mean 65 inches and standard deviation 5 inches.

What is the probability that a randomly selected female student has height between 55 inches and 72.5 inches ?

Example 1

The height of female students entering a university has a normal distribution with mean 65 inches and standard deviation 5 inches. What is the probability that a randomly selected female student has height between 55 inches and 72.5 inches ?

X is the height of a randomly chosen student

X is distributed Normal(65,5)

Need to find $P(55 < X < 72.5)$

Example 1

The height of female students entering a university has a normal distribution with mean 65 inches and standard deviation 5 inches. What is the probability that a randomly selected female student has height between 55 inches and 72.5 inches ?

$P(55 < X < 72.5)$

$P(\frac{55-65}{5} < \frac{X-65}{5} < \frac{72.5-65}{5})$

$P(-2 < Z < 1.5)$

$P(Z < 1.5) - P(Z < -2)$

Example 2

The height of female students entering a university has a normal distribution with mean 65 inches and standard deviation 5 inches.

How tall does a student need to be, so that only 5% of the other students are taller than her?

Example 2

The height of female students entering a university has a normal distribution with mean 65 inches and standard deviation 5 inches. How tall does a student need to be, so that only 5% of the other students are taller than her?

X is the height of a randomly chosen student

X is distributed Normal(65,5)

Need to find c such that $P(X > c) = 0.05$

By definition, c is the 95^{th} percentile of $N(65, 5)$

Example 2

The height of female students entering a university has a normal distribution with mean 65 inches and standard deviation 5 inches. How tall does a student need to be, so that only 5% of the other students are taller than her?

Look in the Z table for the number closest to 0.95.
The table has 0.94950 and 0.95053.

They correspond to z-scores of 1.64 and 1.65.

Since 0.94950 and 0.95053 are about equally close to 0.95, average the z-scores to get 1.645.

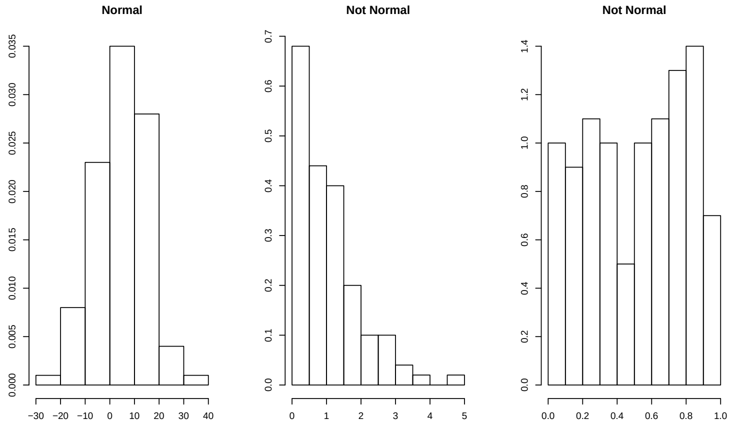
The 95^{th} percentile of $N(65, 5)$ is about
$1.645 \times 5 + 65 = 73.225$.

Assessing Normality

- Sometimes the problem indicates that a r.v. is normally distributed.
- Sometimes we can apply a statistics theorem (coming up next time) to deduce that a r.v. is normally distributed.
- Sometimes we need to assess sample data to figure out if it seems like it came from a normal distribution or not.

Histogram

Check if the sample histogram is symmetric and bell-shaped.

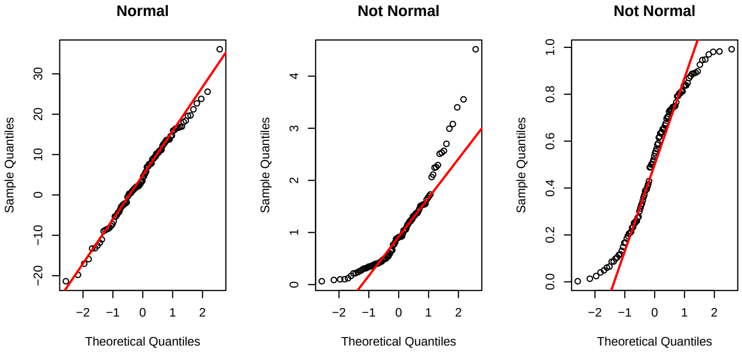


Normal Probability Plot

- y-axis: Observed data, ordered from smallest to largest
- x-axis: Expected z-scores from $N(0, 1)$
- Too complicated to compute by hand
- If data is from a normal distribution, then the points should roughly fall on a line

Normal Probability Plot

Check if the points fall along the line for the most part.



Empirical Rule

Suppose $Z \sim N(0,1)$

- $P(-1 \leq Z \leq 1) \approx 0.68$
- $P(-2 \leq Z \leq 2) \approx 0.95$
- $P(-3 \leq Z \leq 3) \approx 0.997$

Exercise: Verify this using the standard normal table.

Empirical Rule

Suppose $X \sim N(\mu, \sigma)$

- $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68$
- $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95$
- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.997$

Exercise: Verify this using the standard normal table.

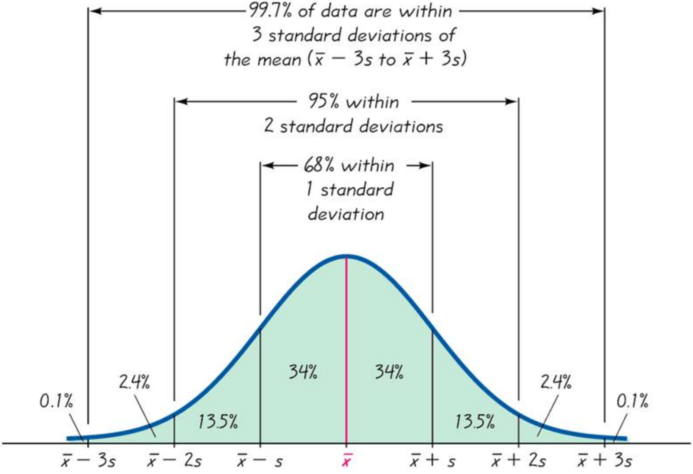
Empirical Rule

Suppose $X \sim N(\mu, \sigma)$

- 68% of observations fall within one standard deviation of the mean.
- 95% of observations fall within two standard deviations of the mean.
- 99.7% of observations fall within three standard deviations of the mean.

Empirical Rule

If a sample comes from a normal distribution, then we should see (roughly) the same proportions.



IQR

- Compute sample IQR and s
- Check if $IQR/s \approx 1.3$
- For a Normal distribution, $IQR/\sigma = 1.34$