STAT 3011 Discussion 015

Week 4

Talha Hamza
University of Minnesota

Spring 2025

Probability Formulas

General Addition Property of Probability:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Independence Rule:
If events A and B are independent, then:

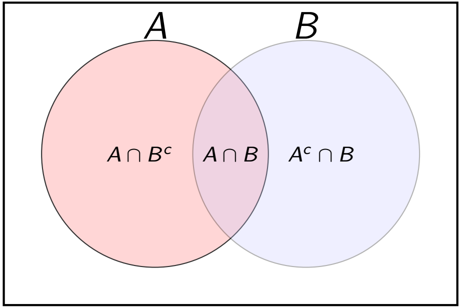
$$P(A \cap B) = P(A) \cdot P(B)$$

Conditional Probability Formula:
Read as Probability of event A occuring GIVEN that B has already occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{where } P(B) > 0$$

Partitioning of Probability

Using a Venn Diagram to Visualize the General Addition Rule:



Partitioning Rule:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

The Complement Rule:

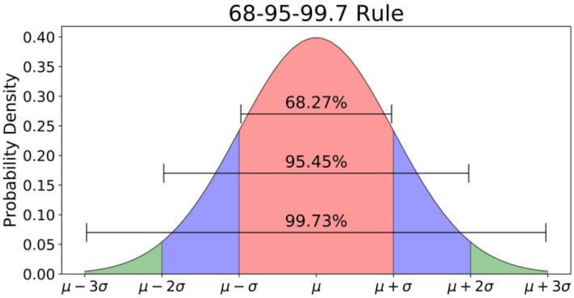
$$P(B^c) = 1 - P(B)$$

The Normal Distribution

- The **normal distribution** is a continuous probability distribution characterized by:
  - **Mean** (μ): The center of the distribution.
  - **Standard deviation** (σ): Measures data spread.
- The **standard normal distribution** is a special case where:
  - $\mu = 0$, $\sigma = 1$
  - Denoted as: $Z \sim N(0, 1)$

The 68-95-99.7 Rule (Empirical Rule)

- This rule describes data distribution in a normal curve:
  - **68%** of data falls within $\pm 1\sigma$.
  - **95%** falls within $\pm 2\sigma$.
  - **99.7%** falls within $\pm 3\sigma$.
- Useful for estimating probabilities and detecting outliers.



Properties of Normal Distribution

Key Characteristics:

- **Bell-shaped curve**: Symmetric and unimodal
- **Mean = Median = Mode**: All measures of center coincide
- **Symmetry**: The distribution is symmetric about the mean μ
- **Asymptotic**: The tails approach but never touch the x-axis
- **Parameter-defined**: Completely determined by μ and σ

Z-Score Transformation

Any normal distribution $X \sim N(\mu, \sigma^2)$ can be converted to the standard normal distribution using the **z-score**:

$$Z = \frac{X - \mu}{\sigma}$$

- This transformation allows us to use standard normal tables or R's `pnorm()` function for probability calculations.
- **Key Idea:** The z-score tells us how far a value is from the mean, making different normal distributions comparable.

Using pnorm() in R

To find the probability under the standard normal curve:

- $P(Z < z)$: **pnorm(z)**
- $P(Z > z)$: **1 - pnorm(z)**
- $P(z_a < Z < z_b)$: **pnorm(z_b) - pnorm(z_a)**

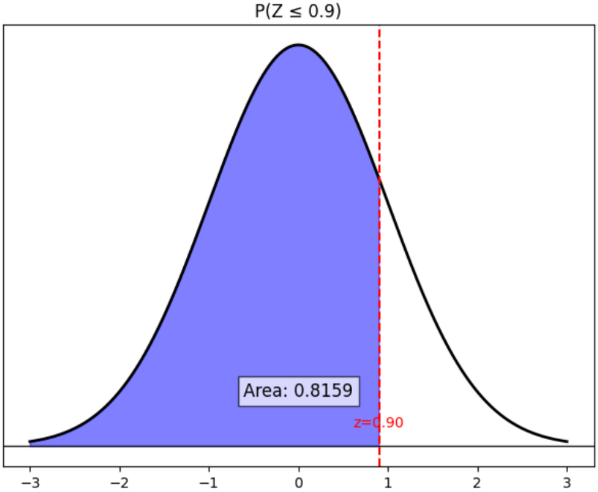
Note: `pnorm()` can also be used for any normal distribution $X \sim N(\mu, \sigma^2)$ by specifying the mean and standard deviation:

- $P(X < x)$: `pnorm(x, mean =` μ`, sd =` σ`)`
- $P(X > x)$: `1 - pnorm(x, mean =` μ`, sd =` σ`)`

Examples

Example 1: Find $P(Z < 0.9)$

```
pnorm(0.9)
# Output: 0.8159
```



Using qnorm() in R

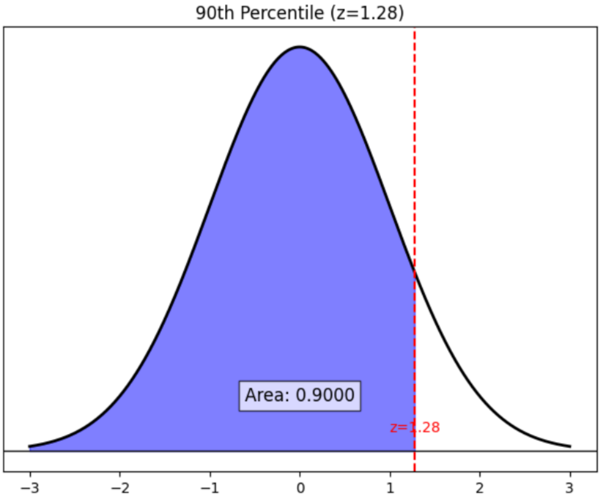
To find the z-value corresponding to a given percentile:

- z such that $P(Z < z) = p$: ‘qnorm(p)‘
- z such that $P(Z > z) = p$: ‘qnorm(1 - p)‘

Example

Example 4: Find the z-value that marks the 90th percentile of the standard normal distribution.

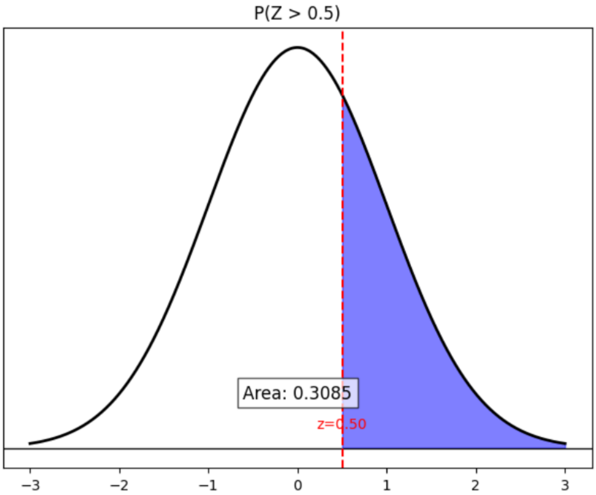
```
qnorm(0.9)
# Output: 1.2816
```



More Examples

Find $P(Z > 0.5)$**.**

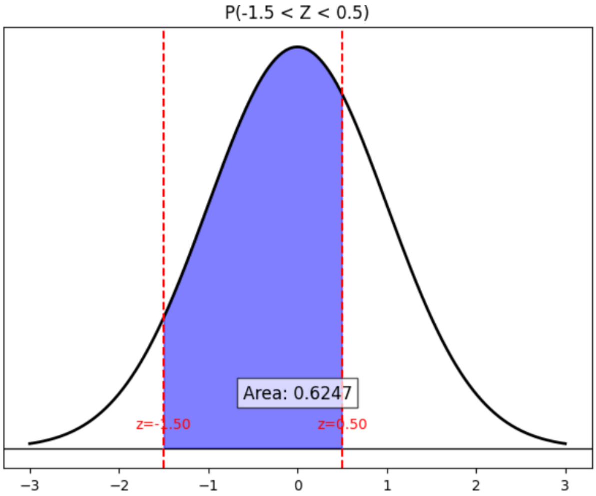
```
1 - pnorm(0.5)
# Output: 0.3085
```



More Examples

Find $P(-1.5 < Z < 0.5)$.

```
pnorm(0.5) - pnorm(-1.5)
# Output: 0.6247
```



pnorm() vs. qnorm()

- **pnorm(z)** $\rightarrow$ Returns the probability (area under the curve) below z.
  "Given a z-score, what proportion of values are smaller?"

- **qnorm(p)** $\rightarrow$ Returns the z-score that corresponds to a given probability.
  "Given a probability, what z-score separates the lower area?"

pnorm() finds the area given a z-score, **qnorm()** finds the z-score given the area.

Exercise 2: Monarch Butterfly Wingspans

Problem Setup: The wingspans of recently cloned monarch butterflies follow a normal distribution with a mean of 9 cm and a standard deviation of 0.75 cm, $X \sim N(9, 0.75)$.

a. Ten percent (10%) of the butterflies have a wingspan narrower than 8.04 cm. Find the 90th percentile of wingspans.
b. Sixty-eight percent (68%) of the butterflies have a wingspan between what two values? Report the answer in cm.
c. What proportion of the butterflies have a wingspan less than 10.5 cm?
d. Is the wingspan 8 cm an unusual value? Show work.
e. Is the wingspan 13 cm an unusual value? Show work.

Solution to Part (a)

Find the 90th percentile of wingspans

- The 10th and 90th percentiles are symmetric about the mean

Solution to Part (a)

Find the 90th percentile of wingspans

- The 10th and 90th percentiles are symmetric about the mean
- Distance between mean and 8.04 cm: $9 - 8.04 = 0.96$ cm

Solution to Part (a)

Find the 90th percentile of wingspans

- The 10th and 90th percentiles are symmetric about the mean
- Distance between mean and 8.04 cm: $9 - 8.04 = 0.96$ cm

Solution to Part (a)

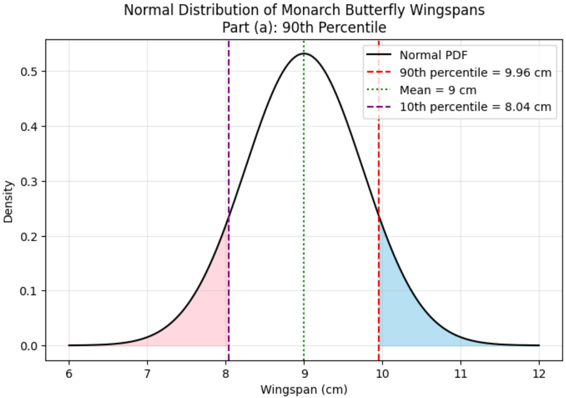
Find the 90th percentile of wingspans

- The 10th and 90th percentiles are symmetric about the mean
- Distance between mean and 8.04 cm: $9 - 8.04 = 0.96$ cm
- Therefore, 90th percentile = mean + 0.96 = $9 + 0.96 = 9.96$ cm

Solution to Part (a)

Find the 90th percentile of wingspans

- The 10th and 90th percentiles are symmetric about the mean
- Distance between mean and 8.04 cm: $9 - 8.04 = 0.96$ cm
- Therefore, 90th percentile $=$ mean $+ 0.96 = 9 + 0.96 = 9.96$ cm



Solution to Part (b)

68% of butterflies have wingspan between what two values?

- Recall the Empirical Rule: 68% of data falls within 1 standard deviation of the mean

Solution to Part (b)

68% of butterflies have wingspan between what two values?

- Recall the Empirical Rule: 68% of data falls within 1 standard deviation of the mean
- Lower bound: $\mu - \sigma = 9 - 0.75 = 8.25$ cm

Solution to Part (b)

68% of butterflies have wingspan between what two values?

- Recall the Empirical Rule: 68% of data falls within 1 standard deviation of the mean
- Lower bound: $\mu - \sigma = 9 - 0.75 = 8.25$ cm
- Upper bound: $\mu + \sigma = 9 + 0.75 = 9.75$ cm

Solution to Part (b)

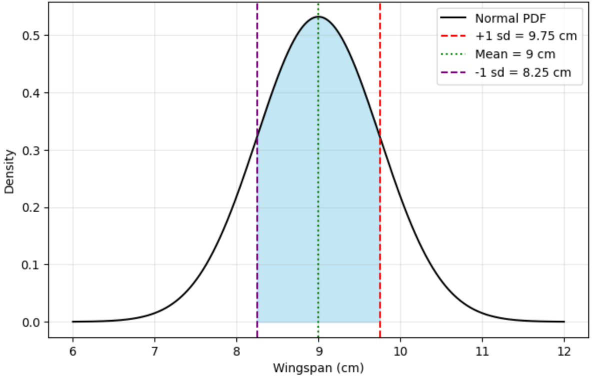
68% of butterflies have wingspan between what two values?

- Recall the Empirical Rule: 68% of data falls within 1 standard deviation of the mean
- Lower bound: $\mu - \sigma = 9 - 0.75 = 8.25$ cm
- Upper bound: $\mu + \sigma = 9 + 0.75 = 9.75$ cm
- Therefore, 68% of butterflies have wingspans between 8.25 cm and 9.75 cm

Solution to Part (b)

68% of butterflies have wingspan between what two values?

- Recall the Empirical Rule: 68% of data falls within 1 standard deviation of the mean
- Lower bound: $\mu - \sigma = 9 - 0.75 = 8.25$ cm
- Upper bound: $\mu + \sigma = 9 + 0.75 = 9.75$ cm
- Therefore, 68% of butterflies have wingspans between 8.25 cm and 9.75 cm



Solution to Part (c)

Proportion with wingspan less than 10.5 cm

Solution to Part (c)

Proportion with wingspan less than 10.5 cm

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{10.5-9}{0.75} = \frac{1.5}{0.75} = 2$

Solution to Part (c)

Proportion with wingspan less than 10.5 cm

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{10.5-9}{0.75} = \frac{1.5}{0.75} = 2$
- Using standard normal distribution: $P(Z < 2) = 0.9772$

Solution to Part (c)

Proportion with wingspan less than 10.5 cm

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{10.5-9}{0.75} = \frac{1.5}{0.75} = 2$
- Using standard normal distribution: $P(Z < 2) = 0.9772$
- Therefore, about 97.72% of butterflies have wingspan less than 10.5 cm

Solution to Part (d)

Is wingspan 8 cm unusual?

Solution to Part (d)

Is wingspan 8 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{8-9}{0.75} = \frac{-1}{0.75} = -1.33$

Solution to Part (d)

Is wingspan 8 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{8-9}{0.75} = \frac{-1}{0.75} = -1.33$
- Check if unusual: Values beyond ± 2 standard deviations are typically considered unusual

Solution to Part (d)

Is wingspan 8 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{8-9}{0.75} = \frac{-1}{0.75} = -1.33$
- Check if unusual: Values beyond ± 2 standard deviations are typically considered unusual
- $|z| = 1.33 < 2$, so it's within 2 standard deviations of the mean

Solution to Part (d)

Is wingspan 8 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{8-9}{0.75} = \frac{-1}{0.75} = -1.33$
- Check if unusual: Values beyond ± 2 standard deviations are typically considered unusual
- $|z| = 1.33 < 2$, so it's within 2 standard deviations of the mean
- **Conclusion:** 8 cm is not an unusual wingspan

Solution to Part (e)

Is wingspan 13 cm unusual?

Solution to Part (e)

Is wingspan 13 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{13-9}{0.75} = \frac{4}{0.75} = 5.33$

Solution to Part (e)

Is wingspan 13 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{13-9}{0.75} = \frac{4}{0.75} = 5.33$
- $|z| = 5.33 > 3$, so it's more than 3 standard deviations from the mean

Solution to Part (e)

Is wingspan 13 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{13-9}{0.75} = \frac{4}{0.75} = 5.33$
- $|z| = 5.33 > 3$, so it's more than 3 standard deviations from the mean
- Using Empirical Rule: 99.7% of data falls within ± 3 standard deviations

Solution to Part (e)

Is wingspan 13 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{13-9}{0.75} = \frac{4}{0.75} = 5.33$
- $|z| = 5.33 > 3$, so it's more than 3 standard deviations from the mean
- Using Empirical Rule: 99.7% of data falls within ± 3 standard deviations
- Only about 0.3% of butterflies would have such extreme wingspans

Solution to Part (e)

Is wingspan 13 cm unusual?

- Calculate z-score: $z = \frac{x-\mu}{\sigma} = \frac{13-9}{0.75} = \frac{4}{0.75} = 5.33$
- $|z| = 5.33 > 3$, so it's more than 3 standard deviations from the mean
- Using Empirical Rule: 99.7% of data falls within ± 3 standard deviations
- Only about 0.3% of butterflies would have such extreme wingspans
- So, 13 cm is a very unusual wingspan

Questions?