STATISTICAL METHODS

Master in Industrial Management,

Operations and Sustainability (MIMOS)

2nd year/1st Semester

2025/2026

CONTACT

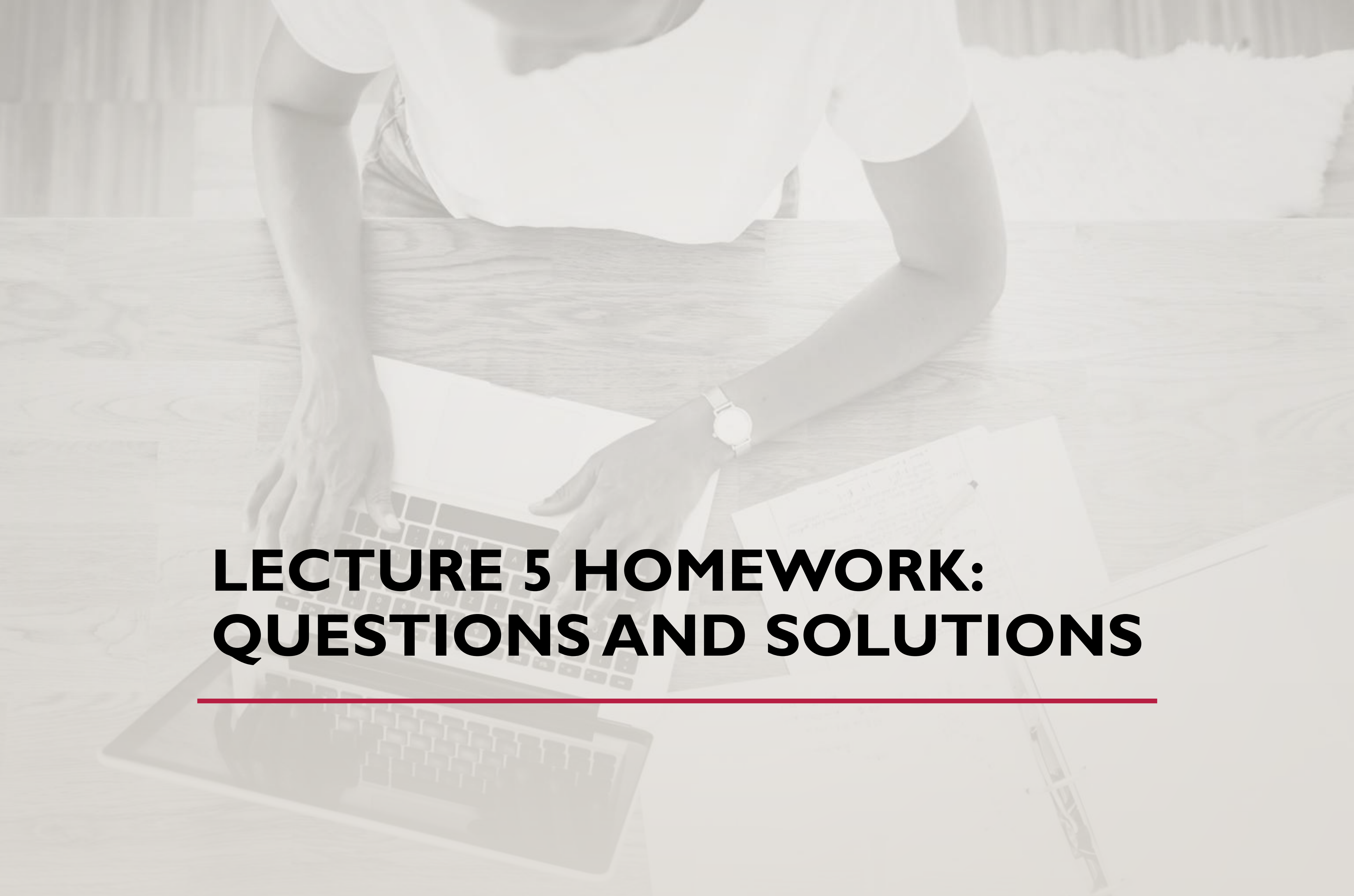
Professor: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

PROGRAM

1. Fundamental Concepts of Statistics
2. Descriptive Data Analysis
3. Introduction to Inferential Analysis
4. Parametric Hypothesis Testing
5. Non-Parametric Hypothesis Testing
6. Linear Regression Analysis



LECTURE 5 HOMEWORK: QUESTIONS AND SOLUTIONS

EXERCISE 4.57

4.57 Records indicate that, on average, 3.2 breakdowns per day occur on an urban highway during the morning rush hour. Assume that the distribution is Poisson.

a. Find the probability that on any given day there will be fewer than 2 breakdowns on this highway during the morning rush hour.
b. Find the probability that on any given day there will be more than 4 breakdowns on this highway during the morning rush hour.

Newbold et al (2013)

EXERCISE 4.57 A): SOLUTION

Answer:

Poisson CDF

Table 6 Cumulative Poisson Probabilities

	Mean Arrival R				
	3.1	3.2	3.3	3.4	3.5
0	.0450	.0408	.0369	.0334	.0302
1	.1847	.1712	.1586	.1468	.1359
2	.4012	.3799	.3594	.3397	.3208
3				.5584	.5366
4				.7442	.7254
5	.9057	.8946	.8829	.8705	.8576
6	.9612	.9554	.9490	.9421	.9347
7	.9858	.9832	.9802	.9769	.9733

Approximate value

$P(X < 2) = P(X \le 1) = F_X(1) \sim 0.1712$

We are given:

$$X \sim Poisson(\lambda = 3.2)$$

The PMF of a Poisson variable:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

a) Probability of fewer than 2 breakdowns

"Fewer than 2" means $X < 2 \Rightarrow X = 0$ or $X = 1$.

$$P(X < 2) = P(X = 0) + P(X = 1)$$

1. $P(X = 0) = \frac{e^{-3.2} 3.2^0}{0!} = e^{-3.2} \approx 0.0408$
2. $P(X = 1) = \frac{e^{-3.2} 3.2^1}{1!} = 3.2 \cdot e^{-3.2} \approx 0.1306$

$$P(X < 2) \approx 0.0408 + 0.1306 = 0.1714$$

✅ So the probability of fewer than 2 breakdowns is approximately **0.171**.

Exact value

EXERCISE 4.57 B): SOLUTION

Answer:

Poisson CDF

Table 6 Cumulative Poisson Probabilities

	Mean Arrival Rate λ			
	3.1	3.2	3.3	3.4
0	.0450	.0408	.0369	.0334
1	.1847	.1712	.1586	.1468
2	.4012	.3799	.3594	.3397
3	.6248	.6025	.5803	.5584
4	.7982	.7806	.7626	.7442
5	.9057	.8946	.8829	.8705
6	.9612	.9554	.9490	.9421
7	.9858	.9832	.9802	.9769
8	[illegible]	[illegible]	[illegible]	[illegible]

$P(X > 4) = 1 - P(X \le 4) =$
$1 - F_X(4) \sim 1 - 0.7806 = 0.2194$

Approximate value

We are given:

$$X \sim Poisson(\lambda = 3.2)$$

The PMF of a Poisson variable:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

b) Probability of more than 4 breakdowns

"More than 4" means $X > 4 \Rightarrow 1 - P(X \le 4)$.

$$P(X > 4) = 1 - P(X \le 4) = 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)]$$

We already have:

- $P(X = 0) \approx 0.0408$
- $P(X = 1) \approx 0.1306$

Compute remaining:

3. $P(X=2) = \frac{3.2^2 e^{-3.2}}{2!} = \frac{10.24}{2} \cdot 0.0408 \approx 0.2080$
4. $P(X=3) = \frac{3.2^3 e^{-3.2}}{3!} = \frac{32.768}{6} \cdot 0.0408 \approx 0.2221$
5. $P(X=4) = \frac{3.2^4 e^{-3.2}}{4!} = \frac{104.8576}{24} \cdot 0.0408 \approx 0.1777$

$$P(X \le 4) = 0.0408 + 0.1306 + 0.2080 + 0.2221 + 0.1777 \approx 0.7792$$

$$P(X > 4) = 1 - 0.7792 \approx 0.2208$$

Exact value

✅ So the probability of more than 4 breakdowns is approximately **0.221**.

EXERCISE 5.6

5.6 The jurisdiction of a rescue team includes emergencies occurring on a stretch of river that is 4 miles long. Experience has shown that the distance along this stretch, measured in miles from its northernmost point, at which an emergency occurs can be represented by a uniformly distributed random variable over the range 0 to 4 miles. Then, if X denotes the distance (in miles) of an emergency from the northernmost point of this stretch of river, its probability density function is as follows:

$$f(x) = \begin{cases} 0.25 & \text{for } 0 < x < 4 \\ 0 & \text{for all other } x \end{cases}$$

a. Graph the probability density function.
b. Find and graph the cumulative distribution function.
c. Find the probability that a given emergency arises within 1 mile of the northernmost point of this stretch of river.
d. The rescue team's base is at the midpoint of this stretch of river. Find the probability that a given emergency arises more than 1.5 miles from this base.

Newbold et al (2013)

EXERCISE 5.6 A): SOLUTION

✅ Answer:

Given:

- $X \sim \text{Uniform}(0, 4)$
- PDF:

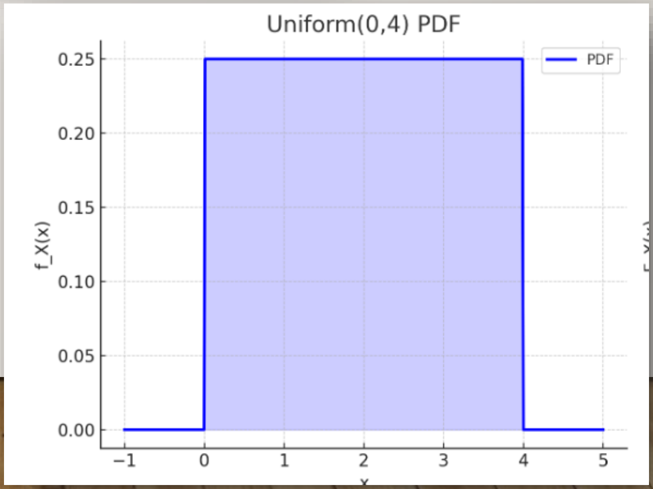
$$f(x) = \begin{cases} 0.25, & 0 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

Check: Uniform distribution has PDF $f(x) = \frac{1}{b-a} = \frac{1}{4-0} = 0.25$. ✅

a. Graph the PDF

- The PDF is constant at 0.25 from $x = 0$ to $x = 4$.
- Outside this range, $f(x) = 0$.

Graph description: A rectangle at height 0.25 from 0 to 4 along the x-axis.



EXERCISE 5.6 B): SOLUTION

✅ Answer:

Given:

- $X \sim \text{Uniform}(0, 4)$
- PDF:

$$f(x) = \begin{cases} 0.25, & 0 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

Check: Uniform distribution has PDF $f(x) = \frac{1}{b-a} = \frac{1}{4-0} = 0.25$. ✅

b. Cumulative Distribution Function (CDF)

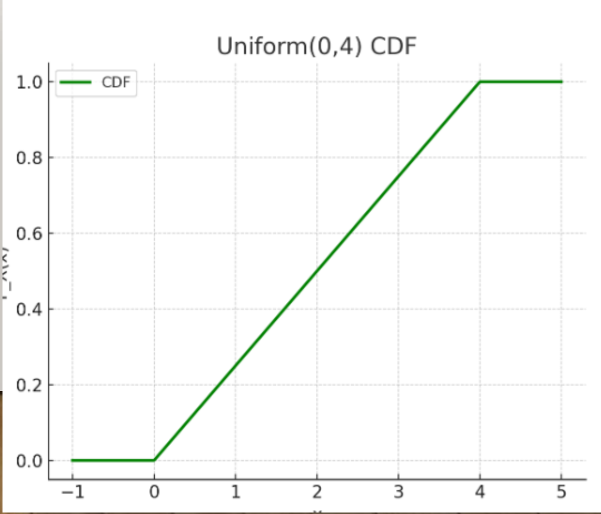
The CDF of a uniform $X \sim U(a, b)$ is:

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

Here, $a = 0, b = 4$, so:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{4} = 0.25x, & 0 \leq x \leq 4 \\ 1, & x > 4 \end{cases}$$

Graph description: A straight line from (0,0) to (4,1), then constant at 1 for $x > 4$.



EXERCISE 5.6 C): SOLUTION

✅ Answer:

Given:

- $X \sim \text{Uniform}(0, 4)$
- PDF:

$$f(x) = \begin{cases} 0.25, & 0 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

Check: Uniform distribution has PDF $f(x) = \frac{1}{b-a} = \frac{1}{4-0} = 0.25$. ✅

c. Probability that an emergency arises within 1 mile of the northernmost point

We want $P(X \leq 1)$:

$$P(X \leq 1) = F(1) = 0.25 \cdot 1 = 0.25$$

✅ So, there is a 25% chance.

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{4} = 0.25x, & 0 \leq x \leq 4 \\ 1, & x > 4 \end{cases}$$

EXERCISE 5.6 D): SOLUTION

Answer:

Given:

- $X \sim \text{Uniform}(0, 4)$
- PDF:

$$f(x) = \begin{cases} 0.25, & 0 \le x \le 4 \\ 0, & \text{otherwise} \end{cases}$$

Check: Uniform distribution has PDF $f(x) = \frac{1}{b-a} = \frac{1}{4-0} = 0.25$. ✅

d. Probability that an emergency arises more than 1.5 miles from the base

- The base is at the **midpoint**: 2 miles.
- "More than 1.5 miles from the base" means $X < 0.5$ **or** $X > 3.5$.

$$P(|X-2| > 1.5) = P(X < 0.5) + P(X > 3.5)$$

- $P(X < 0.5) = F(0.5) = 0.25 \cdot 0.5 = 0.125$
- $P(X > 3.5) = 1 - F(3.5) = 1 - (0.25 \cdot 3.5) = 1 - 0.875 = 0.125$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{4} = 0.25x, & 0 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

$$P(|X-2| > 1.5) = 0.125 + 0.125 = 0.25$$

✅ There is a 25% chance the emergency occurs more than 1.5 miles from the base.

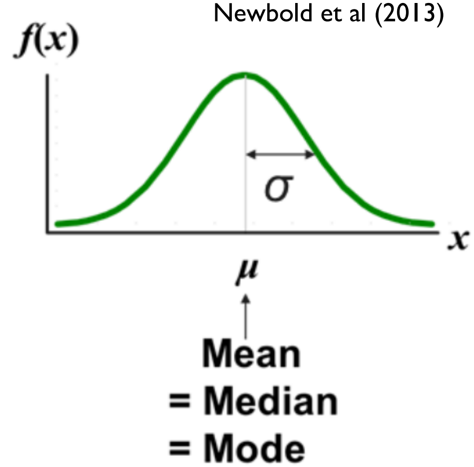
LECTURE 6: NORMAL DISTRIBUTION

NORMAL DISTRIBUTION: CHARACTERISTICS

- Bell Shaped
- Symmetrical
- Mean, Median and Mode are Equal

Probability Density Function (PDF) of a Normal Distribution

Newbold et al (2013)



Location is determined by the mean, μ

Spread is determined by the standard deviation, σ

The random variable has an infinite theoretical range: $+\infty$ to $-\infty$

Mean = Median = Mode, therefore the distribution is symmetric.

- Symmetric distribution
- $E(X) = \boldsymbol{\mu}$
- $Var(X) = \boldsymbol{\sigma^2}$
- Possible values: $]-\infty; +\infty[$

The **normal distribution** depends on the parameters μ and σ^2.
Given the mean μ and variance σ^2 we define the normal distribution using the notation:

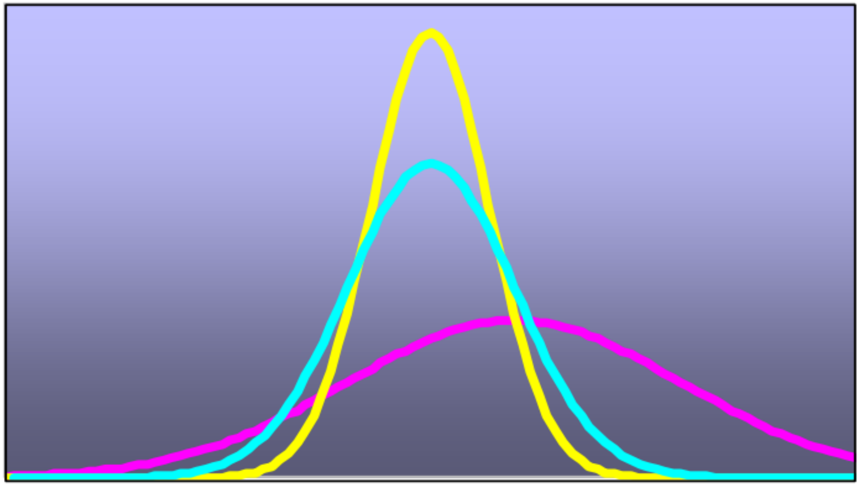
$$X \sim \text{Normal}(\mu, \sigma^2)$$

IMPORTANCE OF THE NORMAL DISTRIBUTION

- The normal distribution closely approximates the probability distributions of a wide range of random variables
- Distributions of sample means approach a normal distribution given a “large” sample size
- Computations of probabilities are direct and elegant
- The normal probability distribution has led to good business decisions for a number of applications

Newbold et al (2013)

MANY NORMAL DISTRIBUTIONS



By varying the parameters μ and σ, we obtain different normal distributions

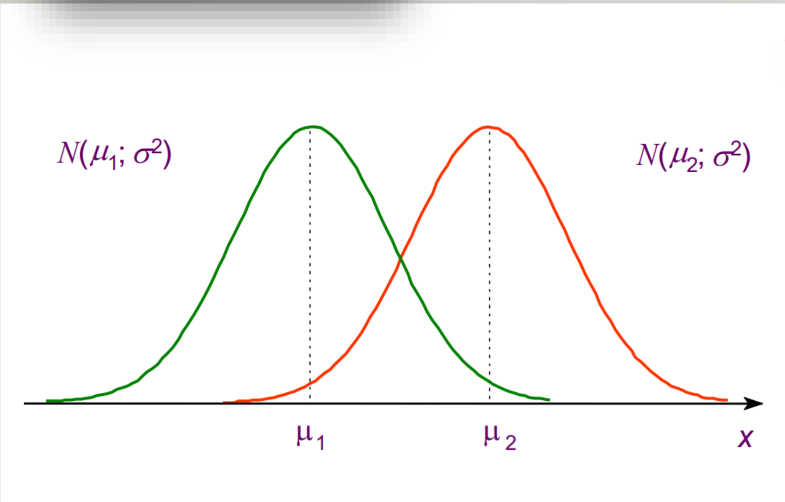
Newbold et al (2013)

NORMAL DISTRIBUTION SHAPE

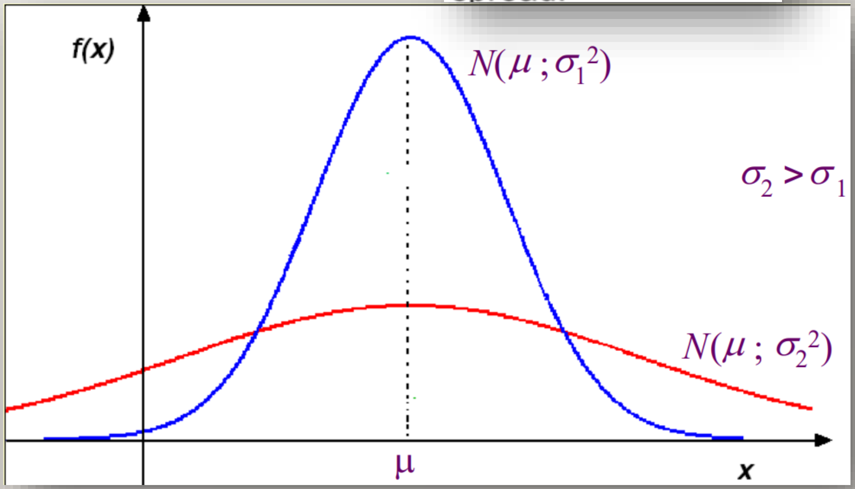
$X \sim \text{Normal}(\mu, \sigma^2)$

The shape of the normal distribution is affected by the two parameters μ and σ².

Changing μ shifts the distribution left or right.



Changing σ increases or decreases the spread.



Normal curves with the same variance σ^2 but different means ($\mu_1 < \mu_2$).

Normal curves with the same mean μ but different variances ($\sigma_1^2 < \sigma_2^2$).

PDF OF A NORMAL DISTRIBUTION

- The formula for the normal probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$X \sim$ Normal(μ, σ^2)

Where e = the mathematical constant approximated by 2.71828

π = the mathematical constant approximated by 3.14159

μ = the population mean

σ^2 = the population variance

x = any value of the continuous variable, $-\infty < x < \infty$

"$e \approx 2.718$, Euler's number"

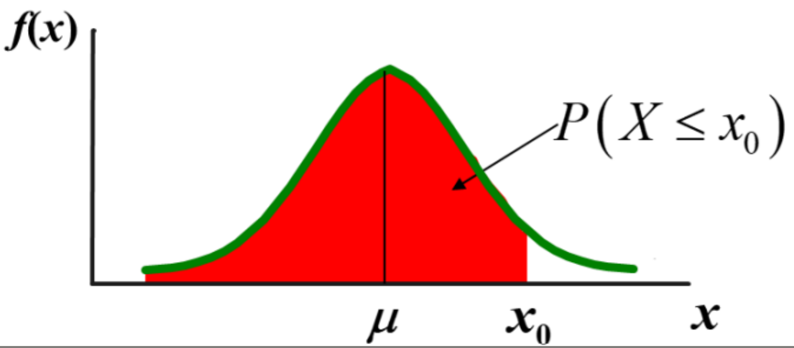
"$\pi \approx 3.1416$, circle constant"

Newbold et al (2013)

CDF OF A NORMAL DISTRIBUTION

- For a normal random variable X with mean μ and variance σ^2, i.e., $X \sim N(\mu, \sigma^2)$, the cumulative distribution function is

$$\boxed{F(x_0) = P(X \le x_0)}$$



The CDF of the normal distribution at x_0 is the area under the curve from $-\infty$ to x_0.

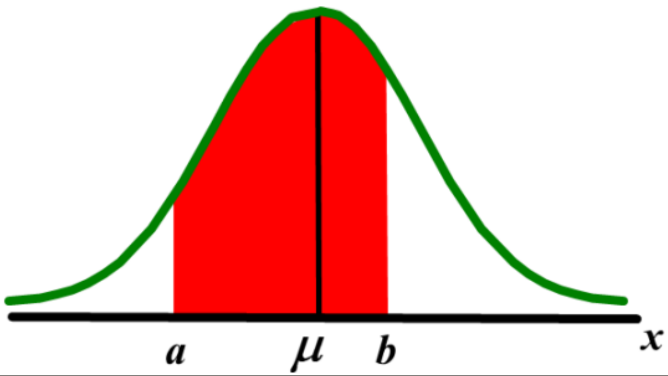
Newbold et al (2013)

FINDING NORMAL PROBABILITIES

The probability for a range of values is measured by the area under the curve

$$P(a < X < b) = F(b) - F(a)$$

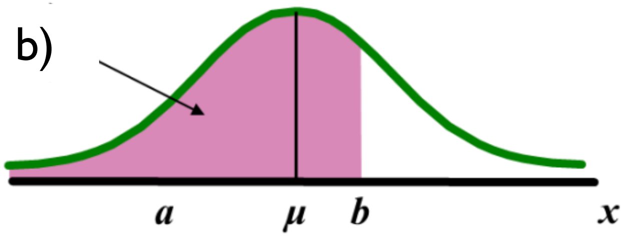
$P(a < X < b)$ is the area under the probability density function (PDF) of the normal distribution between a and b.



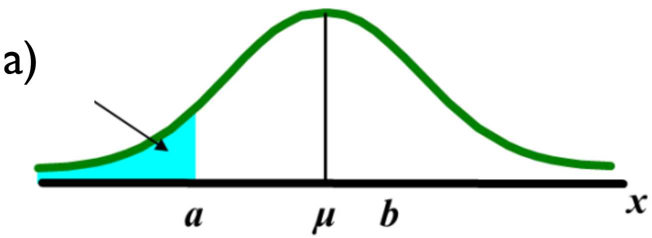
Newbold et al (2013)

FINDING NORMAL PROBABILITIES

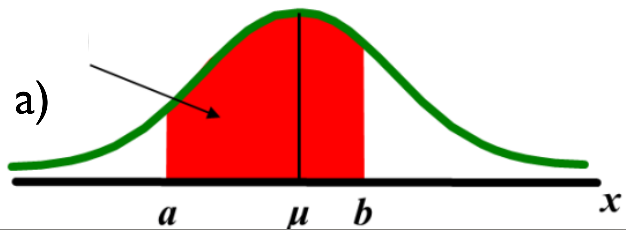
F(b) = P(X < b) = P(X ≤ b)



F(a) = P(X < a) = P(X ≤ a)



F(b) – F(b) = P(a < X < b) =
P(a ≤ X ≤ b) = P(X < b) – P(X< a)

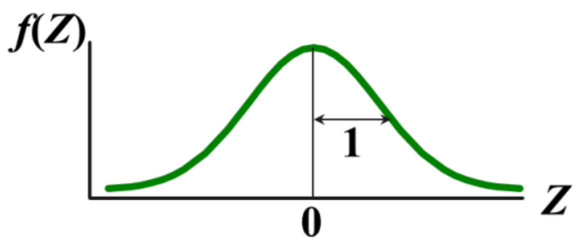


Newbold et al (2013)

STANDARD NORMAL DISTRIBUTION

- Any normal distribution (with any mean and variance combination) can be transformed into the standardized normal distribution (*Z*), with mean 0 and variance 1

$$\boxed{Z \sim N(0,1)}$$



- Need to transform *X* units into *Z* units by subtracting the mean of *X* and dividing by its standard deviation

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

The letter **Z** is reserved for the standard normal distribution.

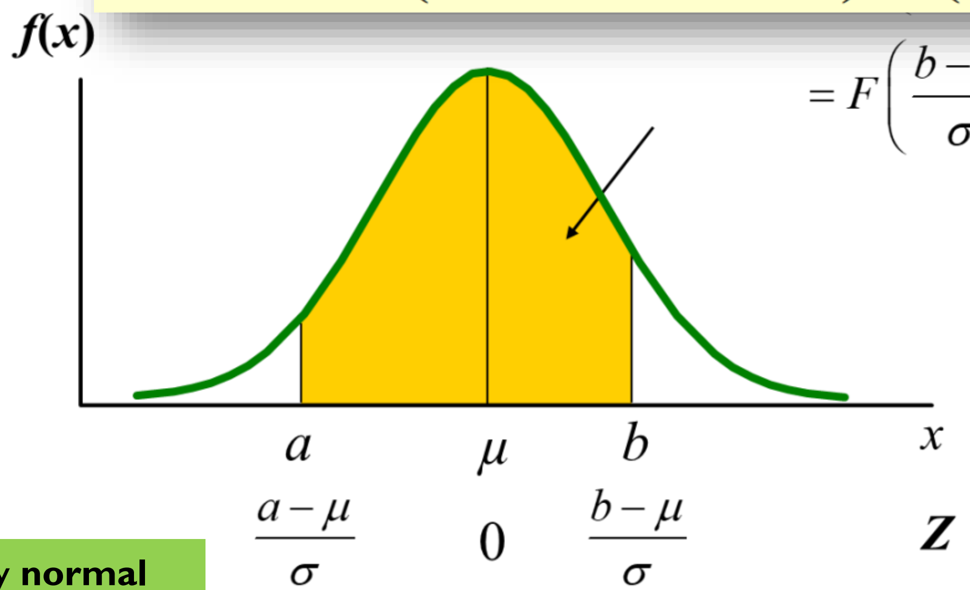
The **standard normal** is a particular form of the normal distribution, with **mean 0** and **standard deviation 1**, and its probabilities are available in tables.

Newbold et al (2013)

NORMAL TO STANDARD NORMAL TRANSFORMATION

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right)$$

$$= F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right)$$



Probabilities for any **normal distribution** are calculated by converting to **Z**, since tables exist only for the standard normal.

Newbold et al (2013)

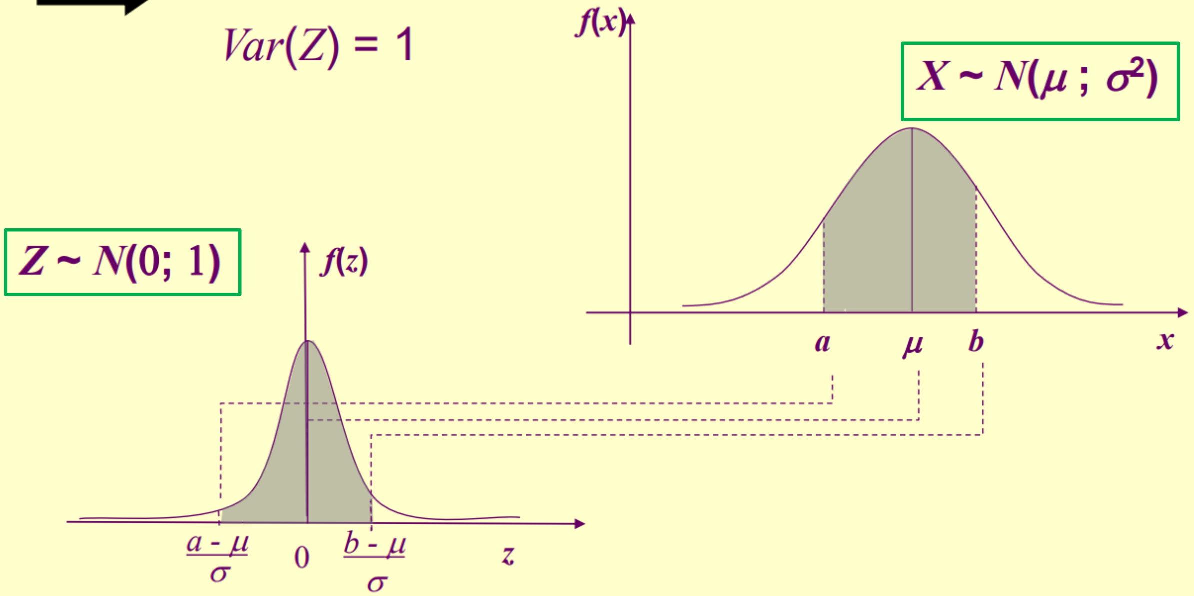
$X \sim$ Normal(μ, σ^2) → **$Z = \frac{X-\mu}{\sigma} \sim$ Normal(0, 1)**

NORMAL TO STANDARD NORMAL TRANSFORMATION

$X \sim \text{Normal}(\mu, \sigma^2)$ → $Z = \dfrac{X - \mu}{\sigma} \sim \text{Normal}(0, 1)$

$E(Z) = 0$

$Var(Z) = 1$



NORMAL TO STANDARD NORMAL TRANSFORMATION: EXAMPLE

- If X is distributed normally with mean of 100 and standard deviation of 50, the Z value for X = 200 is

$$\boxed{Z = \frac{X - \mu}{\sigma} = \frac{200 - 100}{50} = 2.0}$$

$X \sim$ Normal$(100, 50^2)$

$Z = \frac{X - 100}{50} \sim$ Normal$(0, 1)$

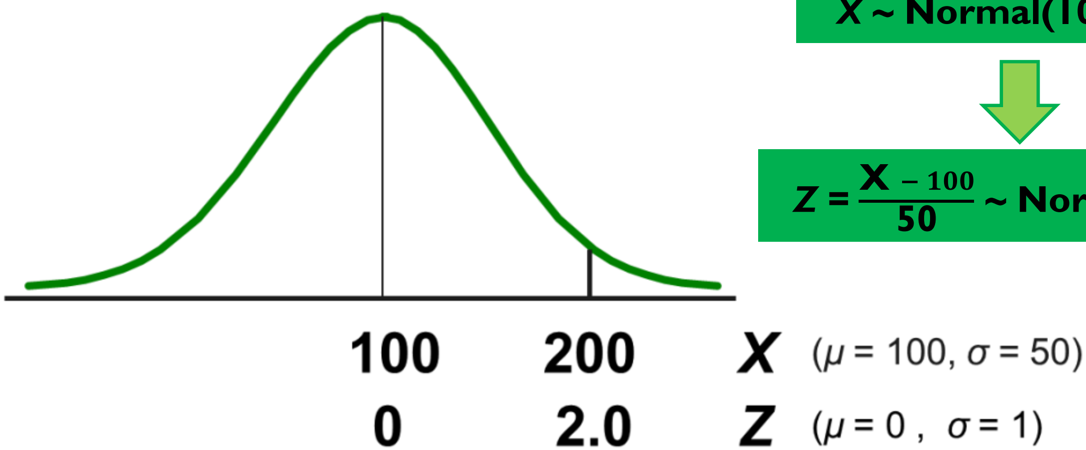
- This says that X = 200 is two standard deviations (2 increments of 50 units) above the mean of 100.

Newbold et al (2013)

NORMAL TO STANDARD NORMAL TRANSFORMATION: EXAMPLE

$X \sim \text{Normal}(100, 50^2)$

$Z = \dfrac{X - 100}{50} \sim \text{Normal}(0, 1)$

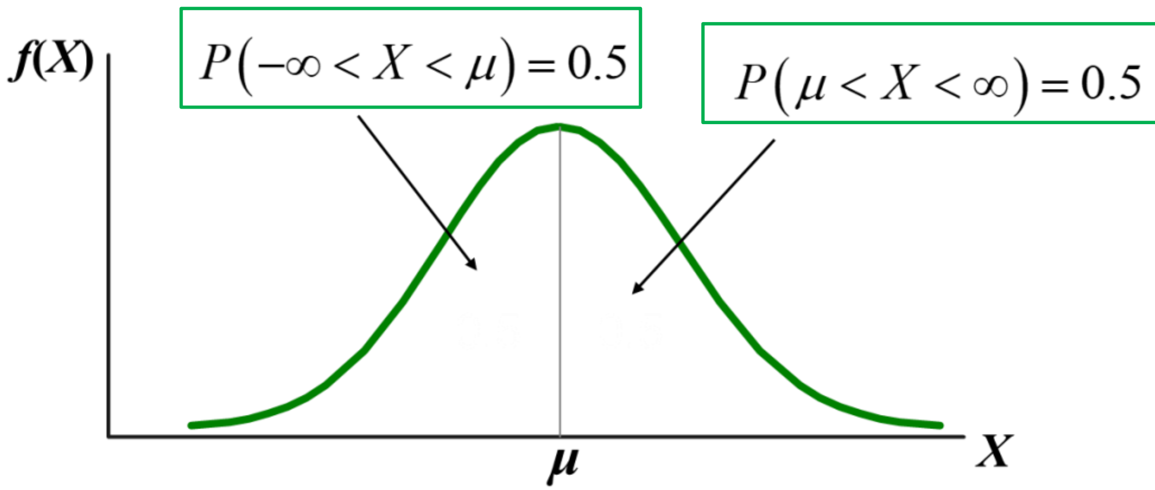


Note that the distribution is the same, only the scale has changed. We can express the problem in original units (X) or in standardized units (Z)

Newbold et al (2013)

PROBABILITY AS AREA UNDER THE CURVE

The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below

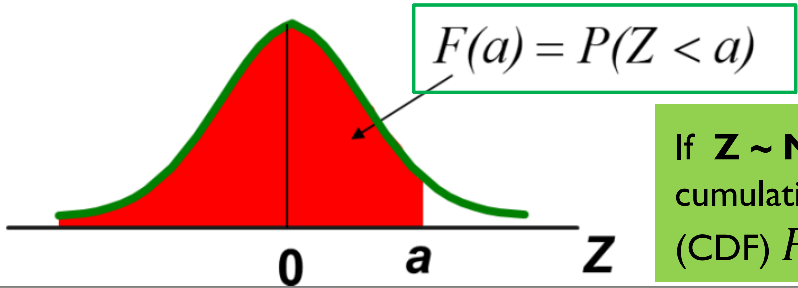


$$P(-\infty < X < \infty) = 1.0$$

Newbold et al (2013)

STANDARD NORMAL TABLE

- The Standard Normal Distribution table in the textbook (Appendix Table 1) shows values of the cumulative normal distribution function
- For a given *Z*-value *a*, the table shows $F(a)$
  (the area under the curve from negative infinity to ***a***)



If **Z ~ N(0, 1)** then the cumulative distribution function (CDF) $F(z)$ equals $\boldsymbol{\Phi(z)}$.

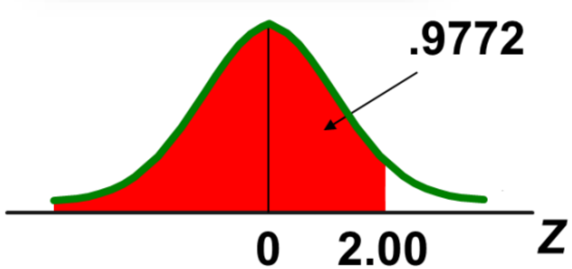
Newbold et al (2013)

STANDARD NORMAL TABLE: EXAMPLE

- Appendix Table 1 gives the probability $F(a)$ for any value a

Example:

$$P(Z < 2.00) = .9772$$



Newbold et al (2013)

Table Cumulative Distribution Function, $F(z)$, of the Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857

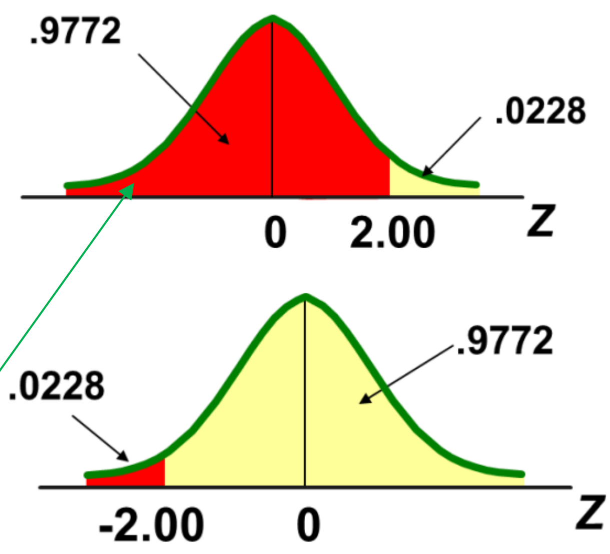
Solution:
$P(Z < 2.00) = F(2.00) = \Phi(2.00) = 0.9772$

STANDARD NORMAL TABLE: EXAMPLE

- For negative Z-values, use the fact that the distribution is symmetric to find the needed probability:

Example:

$$P(Z < -2.00) = 1 - 0.9772$$
$$= 0.0228$$



Newbold et al (2013)

Solution:
P(Z < -2.00) = $\Phi\ (-2.00)$ = 1- $\Phi\ (2.00)$ = 1-0.9772 = 0.0228

GENERAL PROCEDURE FOR FINDING PROBABILITIES

To find $P(a < X < b)$ when X is distributed normally:

- Draw the normal curve for the problem in terms of X
- Translate X-values to Z-values
- Use the Cumulative Normal Table

Newbold et al (2013)

To calculate probabilities for a normal distribution, we must use **a table or a software package;** they cannot be determined directly by hand.

LOWER TAIL PROBABILITIES: EXAMPLE

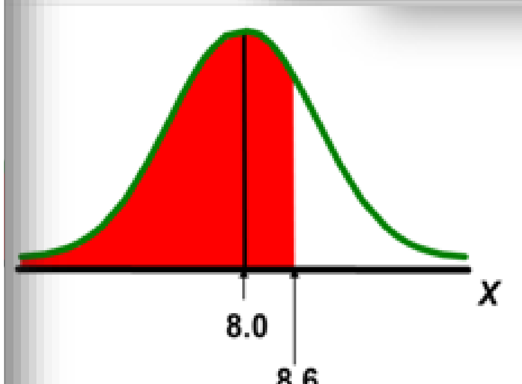
- Suppose *X* is normal with mean 8.0 and standard deviation 5.0

$X \sim$ Normal$(8, 0.5^2)$

Find $P(X < 8.6)$

Table 1 Cumulative Distribution Function, $F(z)$, of the Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	[illegible]	[illegible]	[illegible]	[illegible]
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	[illegible]	[illegible]	[illegible]	[illegible]
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	[illegible]	[illegible]	[illegible]	[illegible]
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	[illegible]	[illegible]	[illegible]	[illegible]
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	[illegible]	[illegible]	[illegible]	[illegible]



Newbold et al (2013)

Solution:

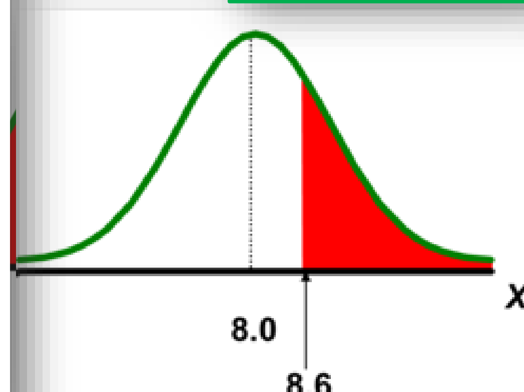
$$P(X < 8.6) = P\left(\frac{X-8}{0.5} < \frac{8.6-8}{0.5}\right) = P(Z < 0.12) = \Phi(0.12) = 0.5478$$

UPPER TAIL PROBABILITIES: EXAMPLE

- Suppose X is normal with mean 8.0 and standard deviation 5.0

$X \sim \text{Normal}(8, 0.5^2)$

$$P(X > 8.6)$$



Newbold et al (2013)

Table 1 Cumulative Distribution Function, $F(z)$, of the Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.8	0.9641	0.9649	0.9656	0.9664	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.9	0.9713	0.9719	0.9726	0.9732	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2.0	0.9772	0.9778	0.9783	0.9788	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2.1	0.9821	0.9826	0.9830	0.9834	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

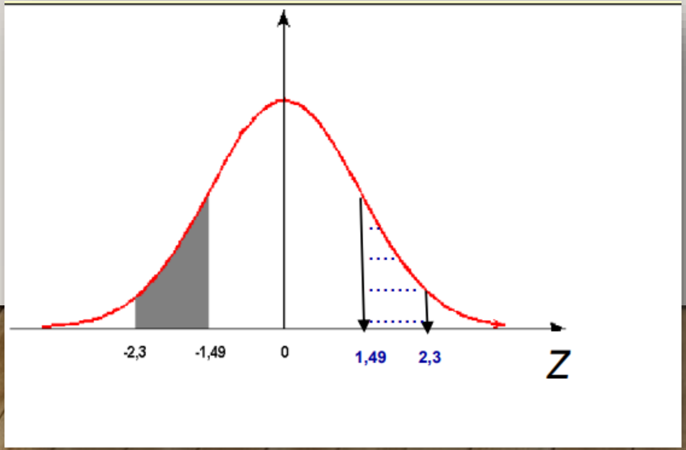
Solution:

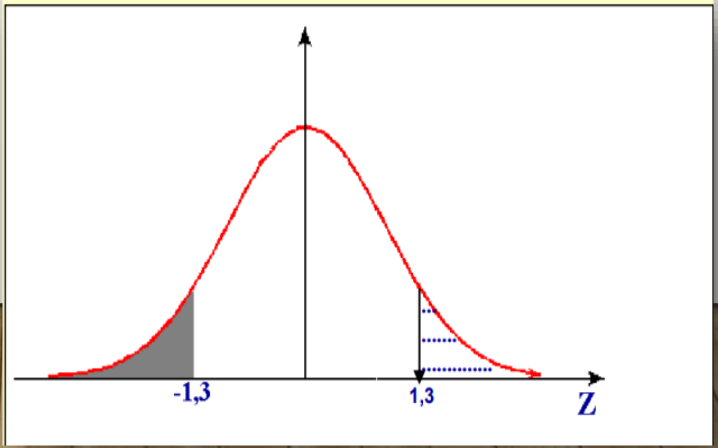
$$P(X > 8.6) = P\left(\frac{X-8}{0.5} > \frac{8.6-8}{0.5}\right) = P(Z > 0.12) = 1 - \Phi(0.12) = 1\text{-}0.5478 = 0.4522$$

PROPERTIES OF THE STANDARD NORMAL CDF

- **$P(-a < Z < a) = \Phi(a) - \Phi(-a)$**
- **$P(Z < -a) = \Phi(-a) = 1 - \Phi(a)$**
- **$P(Z > a) = 1 - \Phi(a)$**
- **$P(Z > -a) = \Phi(a)$**

where *a* is a positive constant and Φ (Phi) is the Cumulative Distribution Function (CDF) of the Standard Normal Distribution.





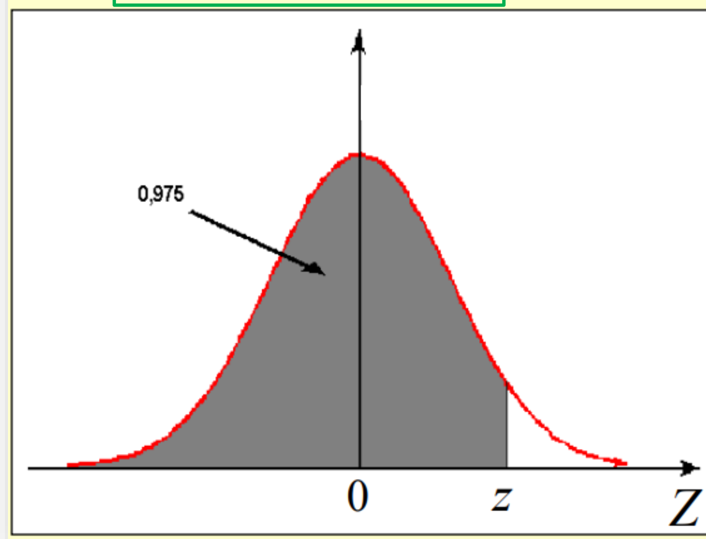
FINDING Z VALUE FOR A GIVEN PROBABILITY

Table 1 Cumulative Distribution Function, $F(z)$, of the Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064		
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443		
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808		
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157		
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486		
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794		
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078		
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340		
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577		
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790		
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980		
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147		
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292		
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418		
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525		
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616		
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693		
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756		
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808		
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	[illegible]	[illegible]

Z ~ Normal(0,1)

(i) $P(Z \leq z) = 0{,}975$



Solution:
$P(Z < z) = 0.975 \Leftrightarrow \Phi(z) = 0.975 \Leftrightarrow z = \Phi^{-1}(0.975) = 1.96$

FINDING X VALUE FOR A GIVEN PROBABILITY

- Steps to find the *X* value for a known probability:
  1. Find the *Z* value for the known probability
  2. Convert to *X* units using the formula:

$$\boxed{X = \mu + Z\sigma}$$

Newbold et al (2013)

If **Z ~ N(0,1)**, we can obtain the variable **X ~ N(μ, σ²)** through the inverse transformation **X = μ + Z×σ.**

FINDING X VALUE FOR A GIVEN PROBABILITY: EXAMPLE

Table 1 Cumulative Distribution Function $F(z)$, of the Standard Normal Distribution Table

z	0	0.01	0.02	0.03	0.04	0.05	0.0
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.52
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.56
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.60
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.64
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.67
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.71
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.74
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.77
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.80
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.83
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.85
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.87
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.89
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.91
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.92
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.94
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.95

Example:

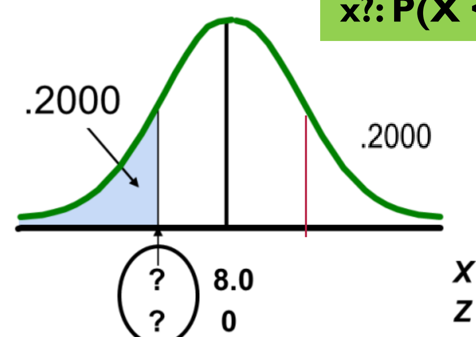
- Suppose X is normal with mean 8.0 and standard deviation 5.0.

$X \sim$ Normal$(8, 0.5^2)$

- Now find the X value so that only 20% of all values are below this X

x?: $P(X < x) = 0.2$

Note that x is the quantile of the normal distribution corresponding to a probability of 0.2, since $P(X < x) = 0.2$.



Solution:

$$P(X < x) = 0.2 \Leftrightarrow P\left(\frac{X-100}{50} < \frac{X-100}{50}\right) = 0.2 \Leftrightarrow$$

$$P\left(Z < \frac{X-100}{50}\right) = 0.2 \Leftrightarrow \Phi\left(\frac{X-100}{50}\right) = 0.2 \Leftrightarrow$$

$$\frac{X-100}{50} = \Phi^{-1}(0.2) \Leftrightarrow \frac{X-100}{50} = -\Phi^{-1}(0.8) \sim -0.84$$

$$\Leftrightarrow \frac{X-100}{50} = -0.84 \Leftrightarrow x = -0.84 \times 50 + 100 = 3.80$$

$\Phi^{-1}(0.2) = z_{0.2}$ is the quantile of the standard normal distribution corresponding to a probability of 0.2. Since this value is negative, it does not appear in the standard normal table, which only lists positive z-values. However, its symmetric value, **$\Phi^{-1}(0.8)$**, is in the table because it is positive. Therefore, we can find $\Phi^{-1}(0.2)$ using $\Phi^{-1}(0.8)$ [$\Phi^{-1}(0.2) = -\Phi^{-1}(1-0.2)$].

EXERCISE 5.23

5.23 Anticipated consumer demand in a restaurant for free-range steaks next month can be modeled by a normal random variable with mean 1,200 pounds and standard deviation 100 pounds.

a. What is the probability that demand will exceed 1,000 pounds?
b. What is the probability that demand will be between 1,100 and 1,300 pounds?
c. The probability is 0.10 that demand will be more than how many pounds?

Newbold et al (2013)

EXERCISE 5.23 A): SOLUTION

✅ Answer:

$$X \sim N(\mu = 1200, \sigma^2 = 100^2)$$

a. Probability that demand exceeds 1,000 pounds

Step 1: Compute z-score

$$Z = \frac{X - \mu}{\sigma} = \frac{1000 - 1200}{100} = -2$$

Step 2: Probability

$$P(X > 1000) = F\,P(Z > -2) \approx 0.9772$$

$$P(Z > -2) = 1 - P(Z \leq -2) = 1 - [1 - P(Z \leq 2)] = P(Z \leq 2)$$

$$P(Z \leq 2) \approx 0.9772 \quad \Rightarrow \quad \boxed{P(Z > -2) \approx 0.9772}$$

✅ Probability ≈ **0.977 (97.7%)**

EXERCISE 5.23 B): SOLUTION

✅ Answer:

$$X \sim N(\mu = 1200, \sigma^2 = 100^2)$$

b. Probability demand is between 1,100 and 1,300

Step 1: Compute z-scores

$$Z_1 = \frac{1100 - 1200}{100} = -1$$

$$Z_2 = \frac{1300 - 1200}{100} = 1$$

Step 2: Probability

$$P(1100 \leq X \leq 1300) = P(-1 \leq Z \leq 1)$$

From z-tables:

$$P(Z < 1) = 0.8413, \quad P(Z < -1) = 0.1587$$

$$P(-1 \leq Z \leq 1) = 0.8413 - 0.1587 = 0.6826$$

✅ Probability ≈ **0.683 (68.3%)**

EXERCISE 5.23 C): SOLUTION

✅ Answer:

$$X \sim N(\mu = 1200, \sigma^2 = 100^2)$$

c. The probability is 0.10 that demand will be more than how many pounds?

We want x such that:

$$P(X > x) = 0.10 \quad \Rightarrow \quad P(X \leq x) = 0.90$$

Step 1: Find z-score for 90th percentile

$$z_{0.90} \approx 1.28$$

Step 2: Convert back to X

$$x = \mu + z\sigma = 1200 + (1.28)(100) = 1200 + 128 = 1328$$

✅ Answer: ≈ **1,328 pounds**

ASSESSING NORMALITY

- Not all continuous random variables are normally distributed
- It is important to evaluate how well the data is approximated by a normal distribution

Newbold et al (2013)

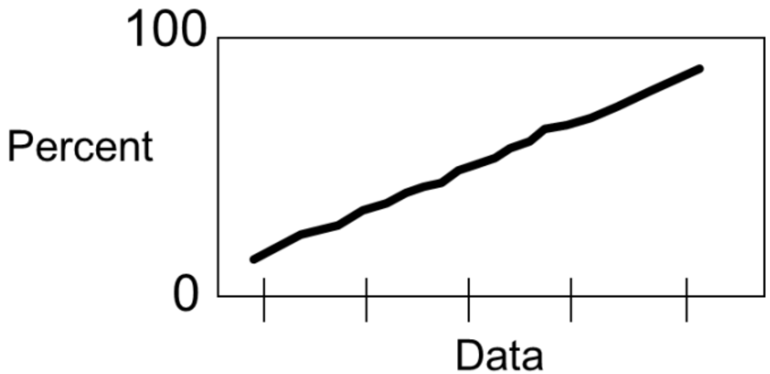
NORMAL PROBABILITY PLOT

- Normal probability plot
  - Arrange data from low to high values
  - Find cumulative normal probabilities for all values
  - Examine a plot of the observed values vs. cumulative probabilities (with the cumulative normal probability on the vertical axis and the observed data values on the horizontal axis)
  - Evaluate the plot for evidence of linearity

Newbold et al (2013)

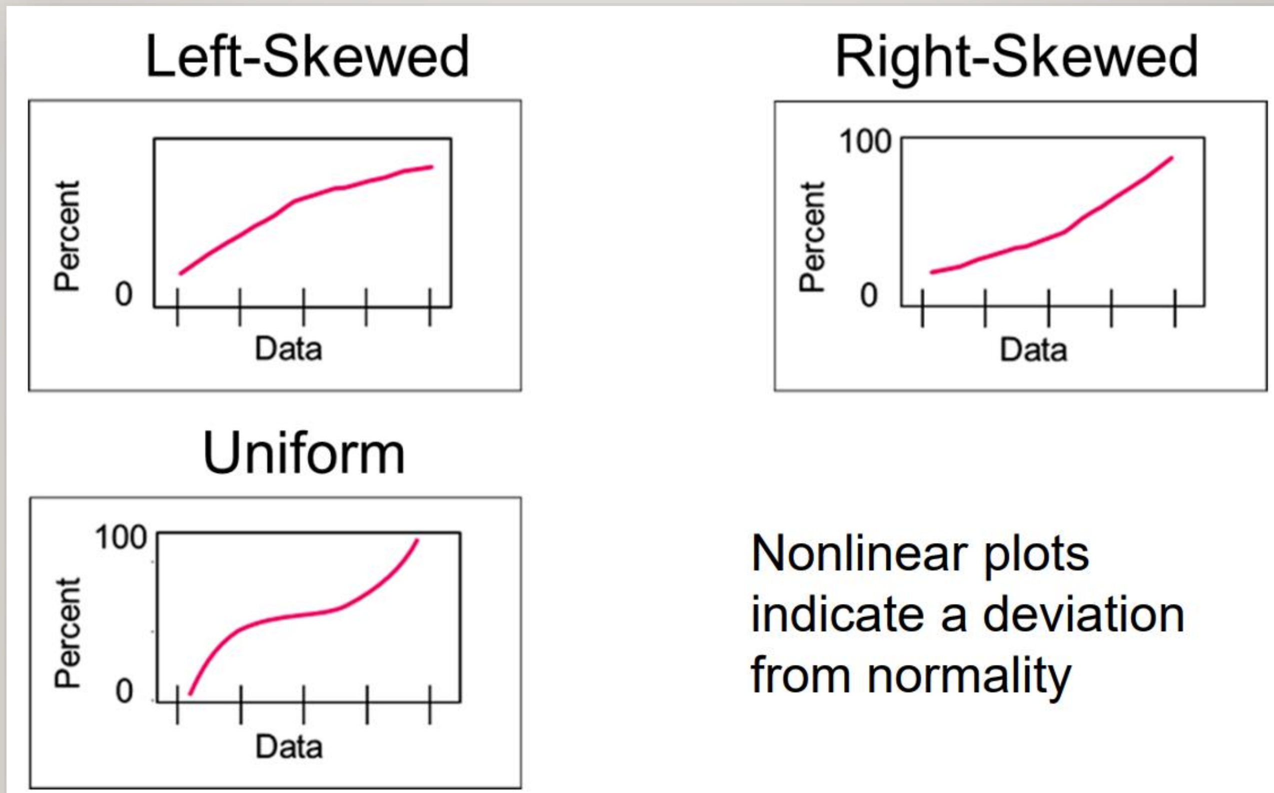
NORMAL PROBABILITY PLOT

A normal probability plot for data from a normal distribution will be approximately linear:



Newbold et al (2013)

NORMAL PROBABILITY PLOT



Newbold et al (2013)

NORMAL DISTRIBUTION APPROXIMATION FOR BINOMIAL DISTRIBUTION

- Recall the binomial distribution:
  - *n* independent trials
  - probability of success on any given trial = *P*
- Random variable *X*:
  - $X_i = 1$ if the *i*th trial is "success"
  - $X_i = 0$ if the *i*th trial is "failure"

X ~ Binomial (n, p)

$$E[X] = \mu = nP$$

$$Var(X) = \sigma^2 = nP(1-P)$$

Newbold et al (2013)

NORMAL DISTRIBUTION APPROXIMATION FOR BINOMIAL DISTRIBUTION

- The shape of the binomial distribution is approximately normal if *n* is large
- The normal is a good approximation to the binomial when $\boxed{nP(1-P) > 5}$
- Standardize to *Z* from a binomial distribution:

$$\boxed{Z = \frac{X - E[X]}{\sqrt{Var(X)}} = \frac{X - np}{\sqrt{nP(1-P)}}}$$

Newbold et al (2013)

$X \sim$ Binomial(n, p) → **$Z = \frac{X - np}{\sqrt{np(1-p)}} \sim$ Normal(0, 1)**

NORMAL DISTRIBUTION APPROXIMATION FOR BINOMIAL DISTRIBUTION

- Let X be the number of successes from n independent trials, each with probability of success P.

- If $nP(1-P) > 5$,

$$\boxed{P(a < X < b) = P\left(\frac{a - nP}{\sqrt{nP(1-P)}} \leq Z \leq \frac{b - nP}{\sqrt{nP(1-P)}}\right)}$$

Newbold et al (2013)

BINOMIAL APPROXIMATION EXAMPLE

- 40% of all voters support ballot proposition *A*. What is the probability that between 76 and 80 voters indicate support in a sample of *n* = 200?
  - $E[X] = \mu = nP = 200(0.40) = 80$
  - $Var(X) = \sigma^2 = nP(1-P) = 200(0.40)(1-0.40) = 48$

    (note : $nP(1-P) = 48 > 5$)

$$\boxed{P(76 < X < 80) = P\left(\frac{76-80}{\sqrt{200(0.4)(1-0.4)}} \le Z \le \frac{80-80}{\sqrt{200(0.4)(1-0.4)}}\right)}$$

$$= P(-0.58 < Z < 0)$$

$$= F(0) - F(-0.58)$$

$$= 0.5000 - 0.2810 = 0.2190$$

We use the standard normal table to calculate these probabilities.

Newbold et al (2013)

NORMAL DISTRIBUTION APPROXIMATION FOR POISSON DISTRIBUTION

Used when the mean (λ) is sufficiently large.

When $\boxed{\lambda \geq 10,}$ the Poisson distribution can be well approximated by a Normal distribution.

In this case:

$$\boxed{X \sim \text{Poisson}(\lambda) \approx N(\mu = \lambda, \sigma^2 = \lambda)}$$

$X \sim$ Poisson(λ) → **$Z = \dfrac{X - \lambda}{\sqrt{\lambda}} \sim$ Normal(0, 1)**

EXERCISE 5.44

5.44 A car-rental company has determined that the probability a car will need service work in any given month is 0.2. The company has 900 cars.

a. What is the probability that more than 200 cars will require service work in a particular month?
b. What is the probability that fewer than 175 cars will need service work in a given month?

Newbold et al (2013)

EXERCISE 5.44: SOLUTION

✅ Answer:

Step 1: Define the random variable

$$X \sim \text{Binomial}(n = 900, p = 0.2)$$

- Mean:

$$\mu = np = 900 \cdot 0.2 = 180$$

- Standard deviation:

$$\sigma = \sqrt{np(1-p)} = \sqrt{900 \cdot 0.2 \cdot 0.8} = \sqrt{144} = 12$$

So $X \sim \text{Binomial}(900, 0.2)$ with $\mu = 180, \sigma = 12$.

Step 2: Apply normal approximation

We approximate with:

$$X \approx N(180, 12^2)$$

Use **continuity correction**.

EXERCISE 5.44 A): SOLUTION

✅ Answer:

a. $P(X > 200)$

With continuity correction:

$$P(X > 200) \approx P(Y > 200.5), \quad Y \sim N(180, 12^2).$$

Compute z-score:

$$z = \frac{200.5 - 180}{12} = \frac{20.5}{12} \approx 1.708$$

From z-tables:

$$P(Z > 1.71) \approx 0.0437$$

✅ Probability ≈ **0.044 (4.4%)**

Continuity Correction for Approximating Binomial with Normal:

When we approximate a **discrete** binomial distribution $X \sim \text{Bin}(n, p)$ using a **continuous** normal distribution $Y \sim N(\mu = np, \sigma^2 = np(1-p))$, we apply a **continuity correction** to improve accuracy.

- **If we want** $P(X \leq k)$: use $P(Y \leq k + 0.5)$
- **If we want** $P(X \geq k)$: use $P(Y \geq k - 0.5)$
- **If we want** $P(X < k)$: use $P(Y \leq k - 0.5)$
- **If we want** $P(X > k)$: use $P(Y \geq k + 0.5)$

💡 In short, **add 0.5 when including the endpoint, subtract 0.5 when excluding it**, depending on the inequality direction.

EXERCISE 5.44 B): SOLUTION

✅ Answer:

b. $P(X < 175)$

With continuity correction:

$$P(X < 175) \approx P(Y < 174.5)$$

Compute z-score:

$$z = \frac{174.5 - 180}{12} = \frac{-5.5}{12} \approx -0.458$$

From z-tables:

$$P(Z < -0.46) \approx 0.322$$

✅ Probability ≈ **0.322 (32.2%)**

Continuity Correction for Approximating Binomial with Normal:

When we approximate a **discrete** binomial distribution $X \sim \text{Bin}(n, p)$ using a **continuous** normal distribution $Y \sim N(\mu = np, \sigma^2 = np(1-p))$, we apply a **continuity correction** to improve accuracy.

- If we want $P(X \le k)$: use $P(Y \le k + 0.5)$
- If we want $P(X \ge k)$: use $P(Y \ge k - 0.5)$
- If we want $P(X < k)$: use $P(Y \le k - 0.5)$
- If we want $P(X > k)$: use $P(Y \ge k + 0.5)$

💡 In short, **add 0.5 when including the endpoint, subtract 0.5 when excluding it**, depending on the inequality direction.

HOMEWORK OF LECTURE 6: QUESTIONS

EXERCISE 5.22

5.22 It is known that amounts of money spent on clothing in a year by students on a particular campus follow a normal distribution with a mean of $380 and a standard deviation of $50.

a. What is the probability that a randomly chosen student will spend less than $400 on clothing in a year?

b. What is the probability that a randomly chosen student will spend more than $360 on clothing in a year?

c. Draw a graph to illustrate why the answers to parts (a) and (b) are the same.

d. What is the probability that a randomly chosen student will spend between $300 and $400 on clothing in a year?

e. Compute a range of yearly clothing expenditures—measured in dollars—that includes 80% of all students on this campus? Explain why any number of such ranges could be found, and find the shortest one.

Newbold et al (2013)

EXERCISE 5.27

5.27 A contractor has concluded from his experience that the cost of building a luxury home is a normally distributed random variable with a mean of \$500,000 and a standard deviation of \$50,000.

a. What is the probability that the cost of building a home will be between \$460,000 and \$540,000?
b. The probability is 0.2 that the cost of building will be less than what amount?
c. Find the shortest range such that the probability is 0.95 that the cost of a luxury home will fall in this range.

Newbold et al (2013)

EXERCISE 5.47

5.47 A hospital finds that 25% of its accounts are at least 1 month in arrears. A random sample of 450 accounts was taken.

a. What is the probability that fewer than 100 accounts in the sample were at least 1 month in arrears?

b. What is the probability that the number of accounts in the sample at least 1 month in arrears was between 120 and 150 (inclusive)?

Newbold et al (2013)

EXERCISE 5.60

5.60 Delivery trucks arrive independently at the Floorstore Regional distribution center with various consumer items from the company's suppliers. The mean number of trucks arriving per hour is 20. Given that a truck has just arrived answer the following:

a. What is the probability that the next truck will not arrive for at least 5 minutes?
b. What is the probability that the next truck will arrive within the next 2 minutes?
c. What is the probability that the next truck will arrive between 4 and 10 minutes?

Newbold et al (2013)

THANKS!

Questions?