2019 – 2020

Dr. Fazıl Küçük Faculty of Medicine, EMU

Year 1 Committee 2

Biostatistics Course

Instructor: Assist. Prof. Dr. İlke Akçay

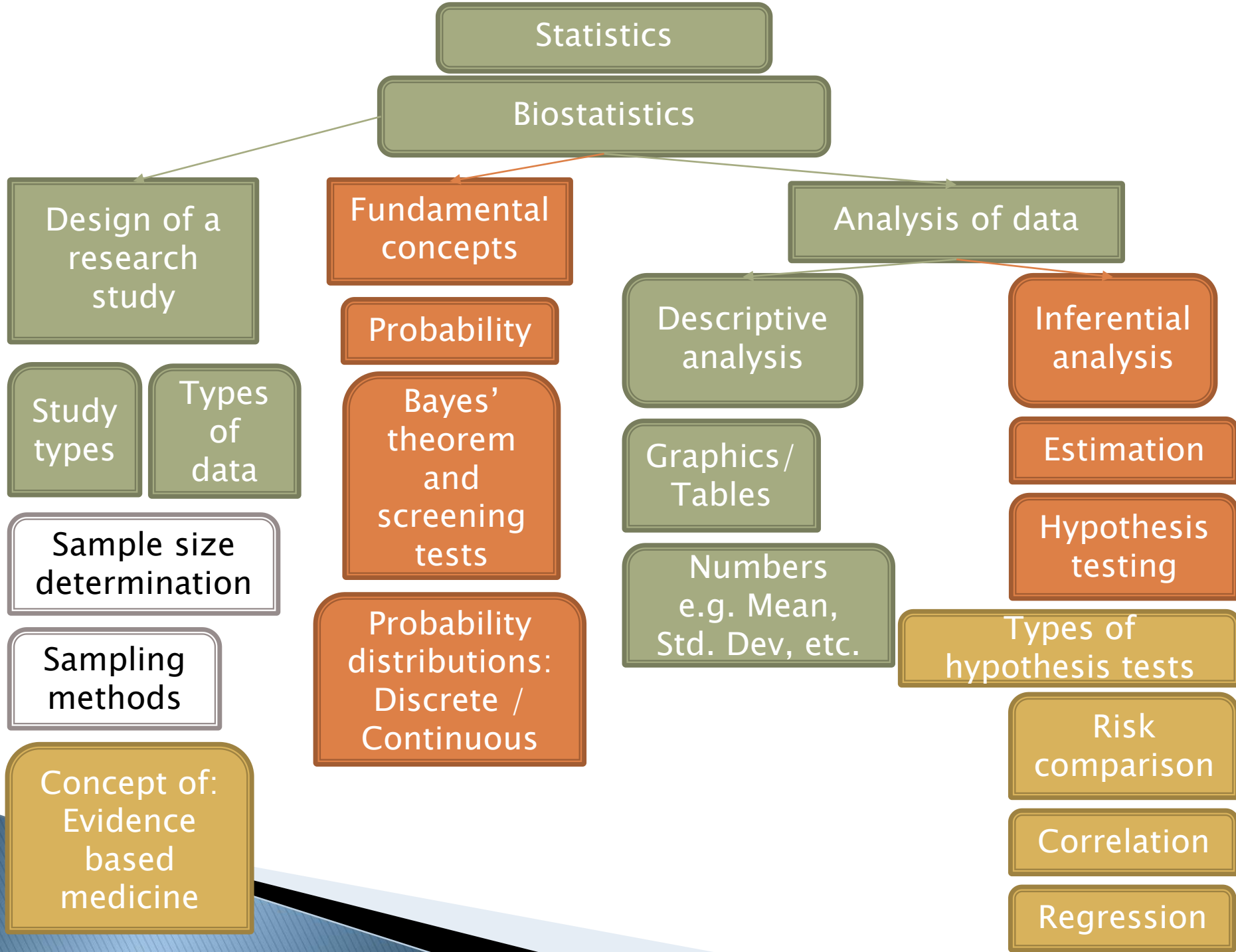
ilke.akcay@emu.edu.tr

Covered lectures in Y1C1

- What is statistics and biostatistics?
- Statistics in medical research
- Designing research (Study types)
- Types of data
- Describing data with graphics
- Describing data with numbers

Topics of Y1C2 biostatistics course

- What is probability and probability distribution?
- Bayes' Theorem
- Principles of statistical analysis
- Elements of statistical inference
- Introduction to statistical analysis
- Sampling, distribution and estimation
- Testing statistical hypothesis
- Types of errors in statistical inference
- Difference between parametric and nonparametric methods; Introduction to parametric methods
- One sample t-test, unpaired t-test and paired t-test


Statistics
Biostatistics
Design of a research study
Study types
Types of data
Sample size determination
Sampling methods
Concept of: Evidence based medicine
Fundamental concepts
Probability
Bayes' theorem and screening tests
Probability distributions: Discrete / Continuous
Analysis of data
Descriptive analysis
Graphics/ Tables
Numbers e.g. Mean, Std. Dev, etc.
Inferential analysis
Estimation
Hypothesis testing
Types of hypothesis tests
Risk comparison
Correlation
Regression

What is probability?

Some Definitions

- **The Universal Set(S)**
  The set of all possible outcomes
- **The empty set** $\emptyset$
  The set containing no elements
- **The event E**
  A set of outcomes in S having a certain characteristic
- **Equally Likely Outcomes**
  The outcomes that have the same chance of occurring
- **Mutually Exclusive (Disjoint) Events**
  Two events are said to be mutually exclusive if they cannot occur simultaneously such that $A \cap B = \emptyset$

Concept of Probability

- **Probability:** the likelihood of a given event's occurrence, which is expressed as a number between 0 and 1
  - The more likely the event, the closer the number is to 1
  - Probability of all possible outcomes of a chance event is always equal to 1

Concept of Probability

- **Subjective Probability**
  - is the probability that measures the confidence that a particular individual has in the truth of a particular proposition.
    - Examples:
    - A physician says that a patient has a 50-50 chance of surviving a certain operation.
    - A researcher suggests that a cure for cancer will be discovered within the next 10 years.

Concept of Probability

- **Objective (Classical) Probability**
  - If an event can occur in N mutually exclusive and equally likely ways, and if m of these possess a trait E, then the probability of the occurrence of event E is equal to m/ N
  - The probability of E is $P(E) = {}^{m}/_{N}$
    - For Example: in the rolling of a die , each of the six sides is equally likely to be observed . So, the probability that a 4 will be observed is equal to 1/6.

Contingency Tables

- Contingency tables (also called as crosstabs or two-way tables) are used in statistics to summarize the relationship between several categorical variables.
- A contingency table is a special type of frequency distribution table, where two variables are shown simultaneously.

- Example: The following is the contingency table related to a medical trial about the effectiveness of a new medication.

Response to medication/ Gender	Female (F)	Male (M)	Total
Positive (Pos)	50	30	80
Negative (Neg)	70	60	130
Total	120	90	210

- Questions:
  - What is the probability that the medicine gives a positive result for females?
    - 50/120
  - What is the probability that the medicine gives a negative result for males?
    - 60/90
  - What is the probability that response to medication is positive?
    - 80/210

Conditional Probability

- $P(A|B)$ is the probability of A assuming that B has happened.
- $P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) \neq 0$ $\quad\quad$ $P(B|A) = \frac{P(A \cap B)}{P(A)}, P(A) \neq 0$

- Example:
- Suppose that a randomly selected subject is a female. What is the probability that her response to medication is positive?
- $P(Pos|F) = \frac{P(Pos \cap F)}{P(F)} = \frac{50/210}{120/210} = \frac{50}{120} = 0.42$

- Suppose we pick a subject at random from the 210 subjects and find that the response to medication is negative. What is the probability that this subject is a female?
- $P(F|Neg) = \frac{P(F \cap Neg)}{P(Neg)} = \frac{70/210}{130/210} = \frac{70}{130} = 0.54$

Joint Probability

- Is calculated when we want to find the probability that a randomly selected subject has two characteristics at the same time
- Example:
- What is the probability that a randomly selected subject is a male and responded to medicine positively?
- $P(M \cap Pos) = \frac{30}{210} = 0.14$

The Multiplication Rule

- $P(A \cap B) = P(A|B)P(B)$
- $P(A \cap B) = P(B|A)P(A)$

- Where,

P(A): marginal probability of A.

P(B): marginal probability of B.

P(B|A):The conditional probability of B given A

- This rule is usually used when the conditional probability and a marginal probability is known.
- Example: Assume that the probability of having breast cancer for a woman in a specific population is 0.21; and the probability that a woman from the same population has cervical cancer given that she has breast cancer is 9%. What is the probability that a randomly selected woman from this population has cervical and breast cancer at the same time?
- $P(B) = 0.21;\ P(C|B) = 0.09$
- $P(C \cap B) = P(C|B)P(B) = 0.21 \times 0.09 = 0.189$

Addition Rule

- For mutually exclusive events A and B, we know that
$$P(A \cup B) = P(A) + P(B)$$
- Example:
- $P(Pos \cup Neg) = P(Pos) + P(Neg) = \frac{80}{210} + \frac{130}{210} = 1$
- What if two events are not mutually exclusive?
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
- Example:
- $P(Pos \cup F) = P(Pos) + P(F) - P(Pos \cap F) = \frac{80+120-50}{210} = \frac{150}{210} = 0.71$

Independent Events

- If two events A and B have no effect on the occurrence of the other, then A and B are called independent events
  - $P(A|B) = P(A)$
  - $P(B|A) = P(B)$
  - $P(A \cap B) = P(B)P(A)$
  - Two events are not independent unless all these statements are true
  - The terms independent and mutually exclusive do not mean the same thing

Complementary Events

- The probability of an event A is equal to 1 minus the probability of its complement, $\bar{A}$

$$P(\bar{A}) = 1 - P(A)$$

- Complementary events are mutually exclusive

- <u>Example:</u>
- $P(Pos) = 1 - P(Neg) = 1 - \frac{130}{210} = \frac{80}{210}$

Marginal Probability

- Given some variable that can be broken down into m categories designated by A1, A2, … , Am and another jointly occurring variable that is broken down into n categories designated by B1, B2, …, Bn; the marginal probability of Ai, P(Ai), equals to the sum of the joint probabilities of Ai with all the categories of B . That is,

$$P(A_i) = \sum P(A_i \cap B_j),$$

for all value of j

Marginal Probability

$$P(A_i) = \sum P(A_i \cap B_j),$$

- Example:
- $P(Pos) = P(Pos \cap F) + P(Pos \cap \mathrm{M}) = \frac{50}{210} + \frac{30}{210} = \frac{80}{210}$

Bayes' Theorem

Bayes' Theorem, Screening Tests, Sensitivity, Specificity, and Predictive Values Positive and Negative

- In health sciences, it is important to correctly predict the presence or absence of a particular disease from knowledge of test results.
- A testing procedure may yield a false positive or false negative
  - A false positive results when a test indicates a positive status when the true status is negative
  - A false negative results when a test indicates a negative status when the true status is positive

Bayes' Theorem, Screening Tests, Sensitivity, Specificity, and Predictive Values Positive and Negative

Test Result	Present (D)	Absent ($\overline{D}$)	Total
Positive (T)	a	b	a+b
Negative ($\overline{T}$)	c	d	c+d
Total	a+c	b+d	n

- The sensitivity of a test is the probability of a positive test result given the presence of the disease is

$$P(T|D) = \frac{a}{a+c}$$

- The specificity of a test is the probability of a negative test result given the absence of the disease is

$$P(\overline{T}|\overline{D}) = \frac{d}{b+d}$$

Bayes' Theorem, Screening Tests, Sensitivity, Specificity, and Predictive Values Positive and Negative

- The predictive value positive of a screening test is the probability that a subject has the disease given that the subject has a positive screening test result: $P(D|T)$

- The predictive value negative of a screening test is the probability that a subject does not have the disease given that the subject has a negative screening test result: $P(\bar{D}|\bar{T})$

Bayes' Theorem, Screening Tests, Sensitivity, Specificity, and Predictive Values Positive and Negative

- Estimates of predictive value positive and predictive value negative of a test may be obtained from knowledge of a test's sensitivity and specificity and the probability of the relevant disease in the general population.
- To obtain the predictive value estimates, we make use of Bayes' Theorem.

Bayes' Theorem, Screening Tests, Sensitivity, Specificity, and Predictive Values Positive and Negative

- Predictive value positive:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|\bar{D})P(\bar{D})}$$

- Predictive value negative

$$P(\bar{D}|\bar{T}) = \frac{P(\bar{T}|\bar{D})P(\bar{D})}{P(\bar{T}|\bar{D})P(\bar{D}) + P(\bar{T}|D)P(D)}$$

Example

- A medical research team wished to evaluate a proposed screening test for Alzheimer's disease on subjects 65 years of age or older. The results are as follows:

Test Result	Yes (D)	No ($\bar{D}$)	Total
Positive(T)	436	5	441
Negative($\bar{T}$)	14	495	509
Total	450	500	950

- In addition, it is estimated that 11.3 percent of the U.S. Population aged 65 years and over have Alzheimer's disease.
  - $P(D) = 0.113$

Example

- Qn1: What is the sensitivity of the test?

$$P(T|D) = \frac{436}{450} = 0.9688$$

- Qn2: What is the specificity of the test?

$$P(\bar{T}|\bar{D}) = \frac{495}{500} = 0.99$$

- Qn3: What is the probability of having a false positive from this test?

$$P(T|\bar{D}) = \frac{P(T \cap \bar{D})}{P(\bar{D})} = \frac{5}{500} = 0.01$$

- Qn4: What is the probability of having a false negative from this test?

$$P(\bar{T}|D) = \frac{P(T \cap D)}{P(D)} = \frac{14}{450} = 0.0311$$

3.5 – Example

- Qn3: What is probability that a subject who is positive on the test has Alzheimer's disease?

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|\bar{D})P(\bar{D})} = \frac{(0.9689)(0.113)}{(0.9689)(0.113) + (0.01)(1 - 0.113)} = 0.93$$

- Qn4: What is probability that a subject who is negative on the test does not have Alzheimer's disease?

$$P(\bar{D}|\bar{T}) = \frac{P(\bar{T}|\bar{D})P(\bar{D})}{P(\bar{T}|\bar{D})P(\bar{D}) + P(\bar{T}|D)P(D)} = \frac{(0.99)(1 - 0.113)}{(0.99)(1 - 0.113) + (0.0311)(0.113)} = 0.996$$

What is a Probability Distribution?

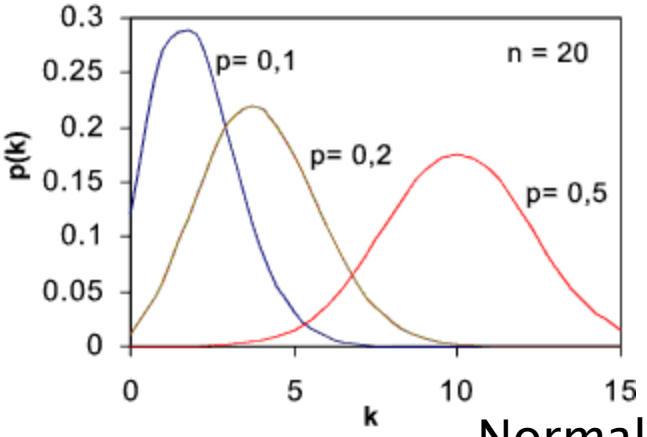
Probability Distribution

- The relationship between the values of a random variable and the probabilities of their occurrence may be summarized by means of a device called a probability distribution.

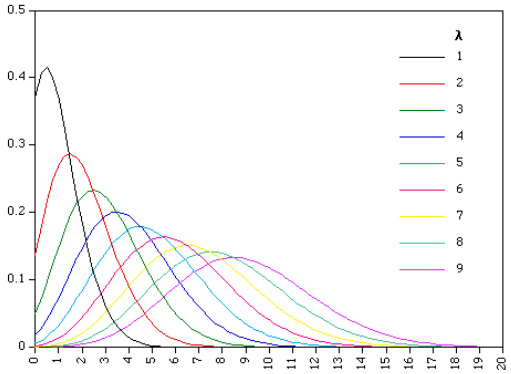
- A probability distribution may be expressed in the form of a table, graph, or formula.

- Knowledge of the probability distribution of a random variable provides the clinician and researcher with a powerful tool
  - For summarizing and describing a set of data, and
  - For reaching conclusions about a population on the basis of a sample drawn from the population.

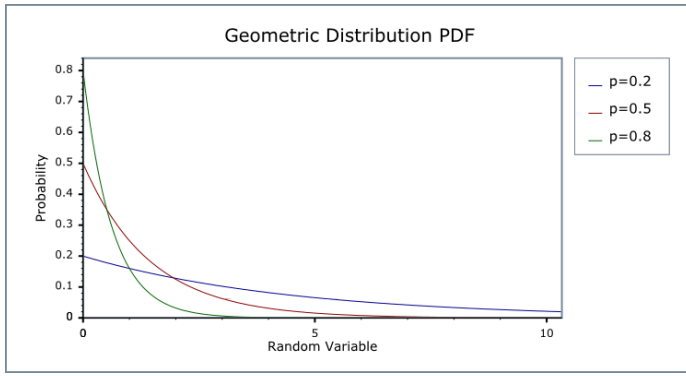
Binomial



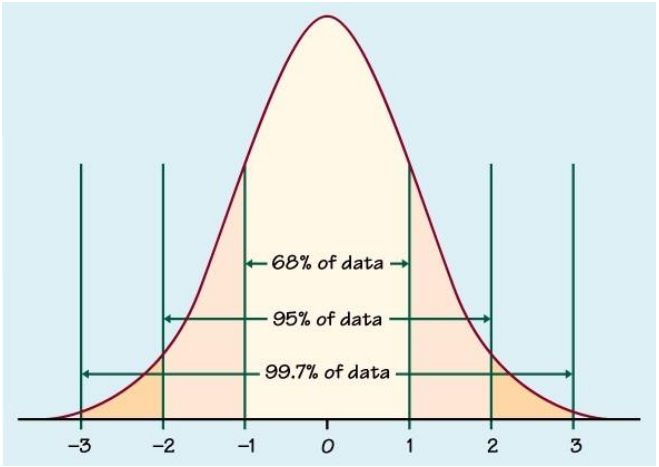
Poisson



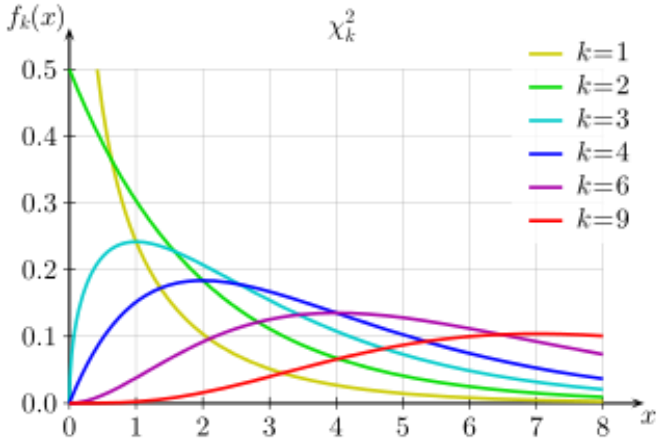
Geometric



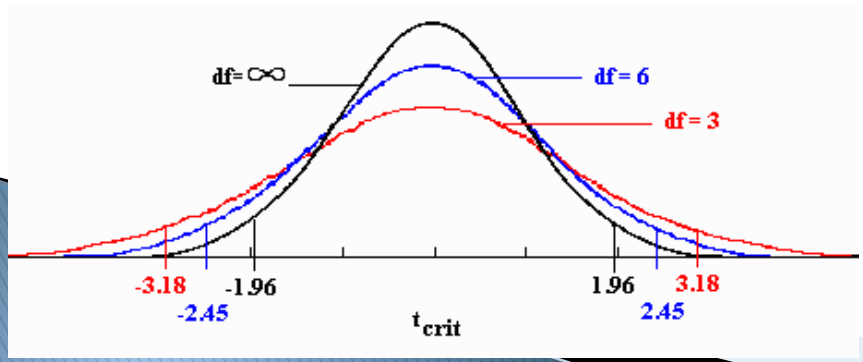
Normal



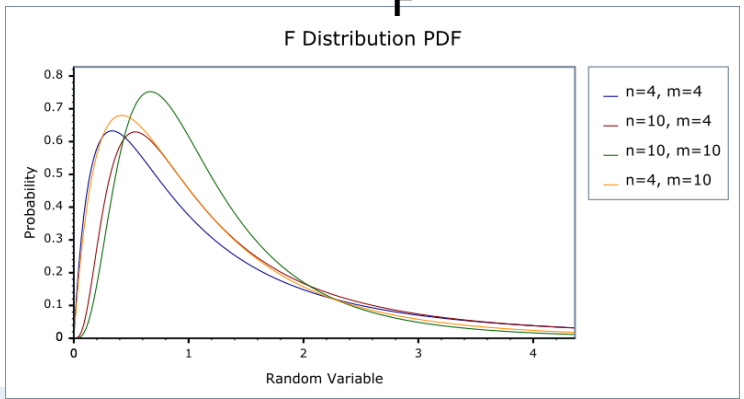
Chi-square



t



F



Discrete Probability Distributions

- Probability Distributions of discrete processes
  - Binomial Distribution
  - Poisson Distribution

Probability Distributions of Discrete Random Variables

- Defn: The Probability Distribution of a **discrete random variable** is a table, graph, formula, or other device used to specify all possible values of a discrete random variable along with their respective probabilities

- It is represented by $f(x) = P(X = x)$

- Properties of probability distribution of discrete random variable:

  $0 \leq P(X = x) \leq 1$

  $\sum P(X = x) = 1$

Example

- Experiment: Two coins are tossed. Let Y equals the number of heads observed.

Possible Outcomes	# of heads (Y)
H, H	2
H, T	1
T, H	1
T, T	0

- So, Y can take on values 0, 1, 2
- The probability distribution for Y is:

Y	P(Y=y)
0	¼
1	½
2	¼
	Total = 1

Probability Distributions of Discrete Random Variables

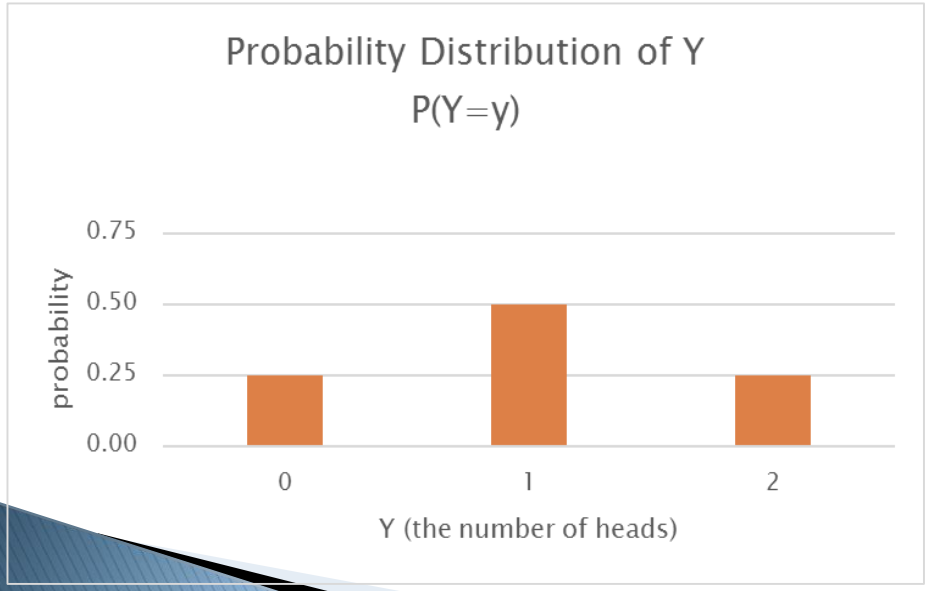
- Defn: The Cumulative Probability Distribution of discrete random variable X shows the probability that X is less than or equal to a certain value x

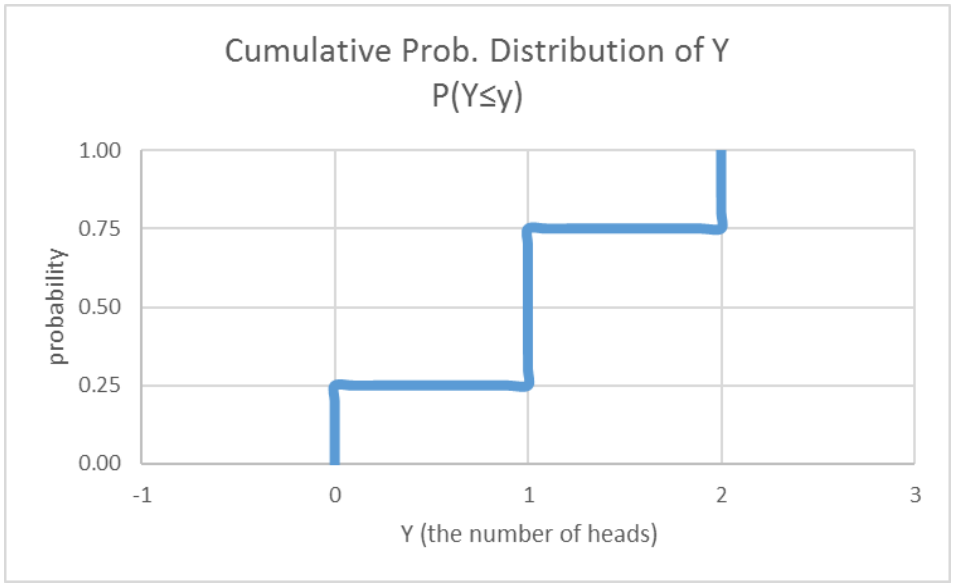
- It is denoted by $F(x) = P(X \leq x)$

Y	P(Y=y)	P(Y≤y)
0	¼	¼
1	½	¾
2	¼	1
	Total = 1	

Graphical Representation of a discrete probability distribution

- the graphical representation of $P(X = x)$
- the graphical representation of $P(X \leq x)$





The Binomial Distribution

- It is derived from a process known as Bernoulli Process
  - 1) The process consists of **n identical trials**
  - 2) Each trial results in one of two possible, mutually exclusive outcomes: **success or failure**
    - dead/alive – sick/well
  - 3) The **probability of success**, denoted by p, remains the same from trial to trial. The probability of failure is $q = 1 - p$
  - 4) The trials are independent, that is, the outcome of any particular trial is not affected by the outcome of any other trial

The Binomial Distribution

- Defn: **The probability distribution of the binomial random variable X**, the number of successes in **n** independent trials is:

- $f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2, \dots, n$

where
- n : number of trials
- x : number of successes
- $n - x$: number of failures
- p : probability of success
- q : probability of failure

Example

- Suppose it is known that 10% of a certain population is color blind. If a random sample of 25 people is drawn from this population, find the probability that
  - a) Exactly 4 of them will be color blind
    - $P(X = 4) = 25C_4(0.1)^4(0.9)^{21} = 0.1384$
  - b) 2 or fewer will be color blind
    - $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

    $= {}_{25}C_0(0.1)^0(0.9)^{25} + {}_{25}C_1(0.1)^1(0.9)^{24} + {}_{25}C_2(0.1)^2(0.9)^{23}$

    $= 0.5371$

Example cont'd

- c) Five or fewer will be color blind
  - $P(X \leq 5)$
- d) Six or more will be color blind
  - $P(X \geq 6)$
- e) Between 6 and 9 inclusive are color blind
  - $P(6 \leq X \leq 9)$
- f) two, three or four will be color blind
  - $P(X = 2) + P(X = 3) + P(X = 4)$

The Poisson Distribution

- In Poisson Distribution, the discrete random variable X is the number of occurences of some random event in a certain period of time or space

- The probability distribution of X is given by:

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

Where e: constant=2.7183

λ: the average number of occurences of the random event in the interval

Example

- Suppose we know that births in a hospital occur randomly at an average rate of 1.8 births per hour. What is the probability that
- a) Exactly 2 births occur in an hour?
  - $x = 2;\ \lambda = 1.8$
  - $P(X = 2) = \frac{e^{-1.8}1.8^2}{2!} = 0.2678$
- b) more than 2 births occur in an hour?
  - $P(X > 2) = 1 - P(X \leq 2)$
  - $= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$
  - $= 1[\frac{e^{-1.8}1.8^0}{0!} + \frac{e^{-1.8}1.8^1}{1!} + \frac{e^{-1.8}1.8^2}{2!}]$

    $= 1 - [0.1653 + 0.2975 + 0.2678] = 0.2694$

Example cont'd

- c) no births occur in a given 2 hour interval?
  - $\lambda = (1.8)(2) = 3.6$
  - $P(X = 0)$
- d) exactly 5 births occur in 2 hour interval?
  - $P(X = 5)$ with $\lambda = (1.8)(2) = 3.6$
- e) At least one birth occurs in a 15-min interval?
  - $P(X \geq 1)$ with $\lambda = (1.8)(1/4) = 0.45$

Continuous Probability Distributions

- Probability Distributions of continuous processes
  - Normal Distribution
  - t Distribution
  - Many others like Chi-square distribution

Continuous Probability Distributions

- Between any two values assumed by a continuous random variable, there exist an infinite number of values
- Smooth curves are used to graphically represent the distributions of continuous random variables
- If n (number of values) is large and the width of the class intervals are small, then we have the following graphical representations

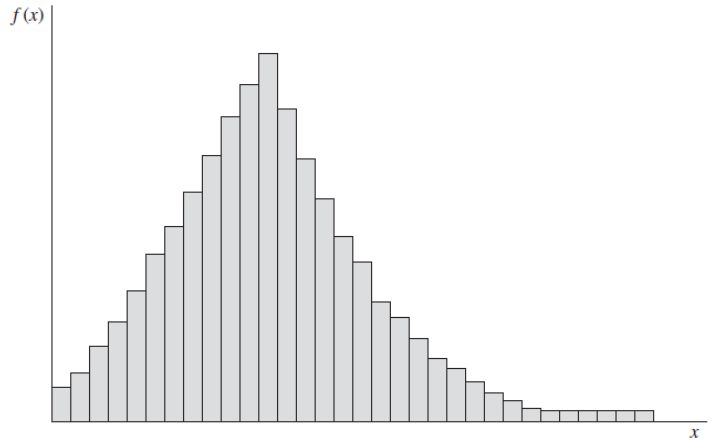


FIGURE 4.5.1 A histogram resulting from a large number of values and small class intervals.

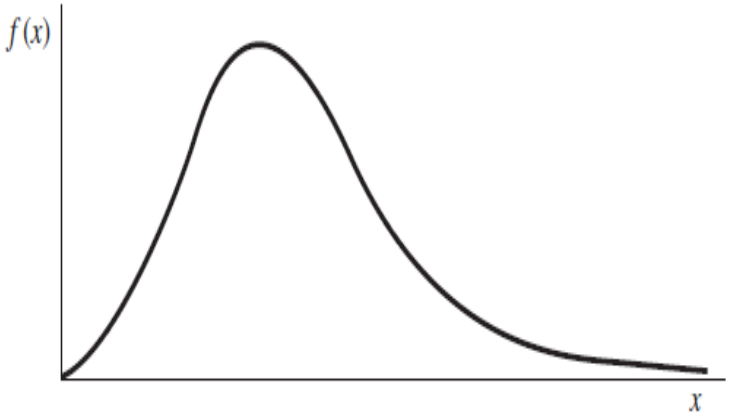


FIGURE 4.5.2 Graphical representation of a continuous distribution.

Properties of continuous probability distributions

- 1) Area under the curve $=1$
  - $\int_{-\infty}^{\infty} f(x) = 1$
- 2) $P(X = a) = 0$ where a is a constant
  - $\int_{a}^{a} f(x) = 0$
- 3) The area between points a,b $= P(a < x < b)$
  - $\int_{a}^{b} f(x)$

The Normal Distribution

- The most important distribution in statistics
- The Normal Density is given by:

- $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$

- π, e : constants
- μ : population mean.
- σ : Population standard deviation

The Normal Distribution

- The graph of the Normal Distribution produces the familiar bell-shaped curve

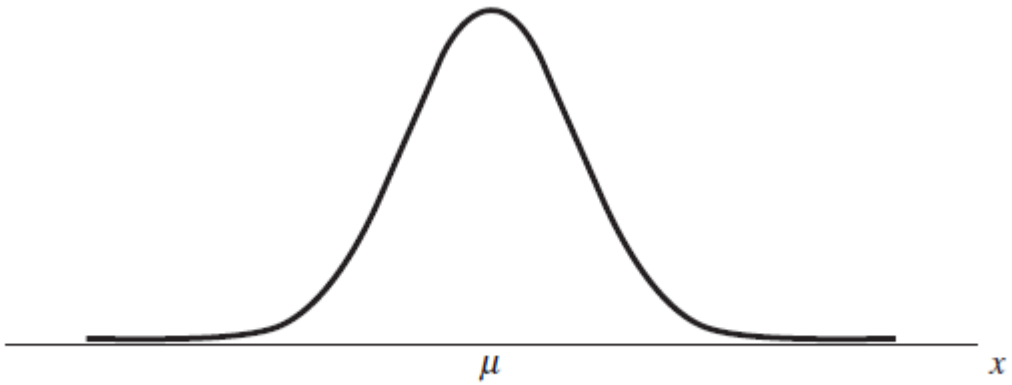


FIGURE 4.6.1 Graph of a normal distribution.

Characteristics of Normal Distribution

- It is symmetrical about its mean
- The mean, median and mode are all equal
- The total area under the curve above the x-axis is 1

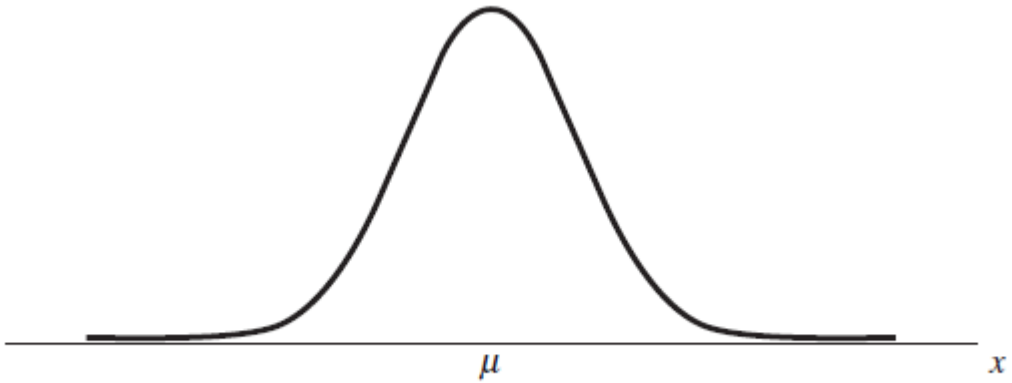
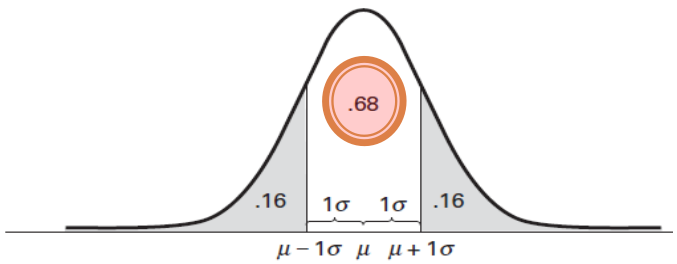
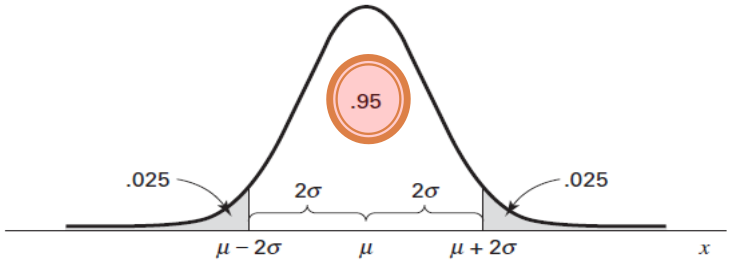


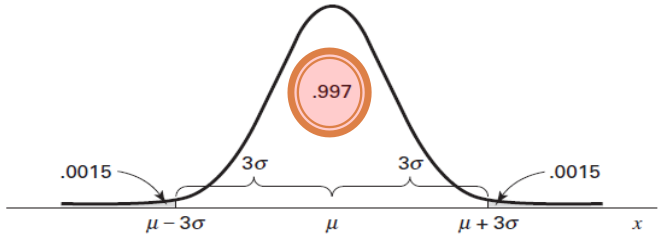
FIGURE 4.6.1 Graph of a normal distribution.

Characteristics of Normal Distribution

- 68% of a normally distributed data is within 1 standard deviation from the mean.
- 95% of a normally distributed data is within 2 standard deviations from the mean.
- 99.7% of a normally distributed data is within 3 standard deviations from the mean.







Characteristics of Normal Distribution

- The Normal Distribution is completely determined by the parameters μ and σ. That is, a different normal distribution is specified for each different value of μ and σ.
- Different values of μ shift the curve along the x-axis
  - -> μ: location parameter

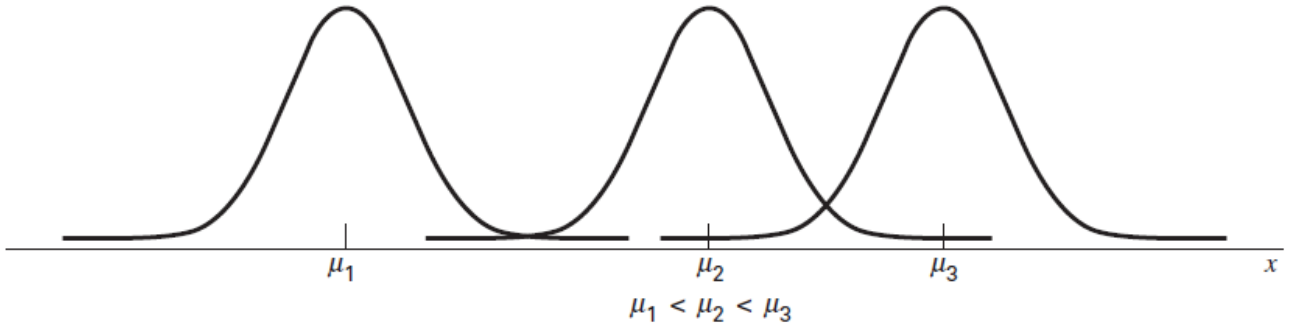


FIGURE 4.6.3 Three normal distributions with different means but the same amount of variability.

Characteristics of Normal Distribution

- Different values of σ determine the degree of flatness or peakedness of the curve
  - -> σ: shape parameter

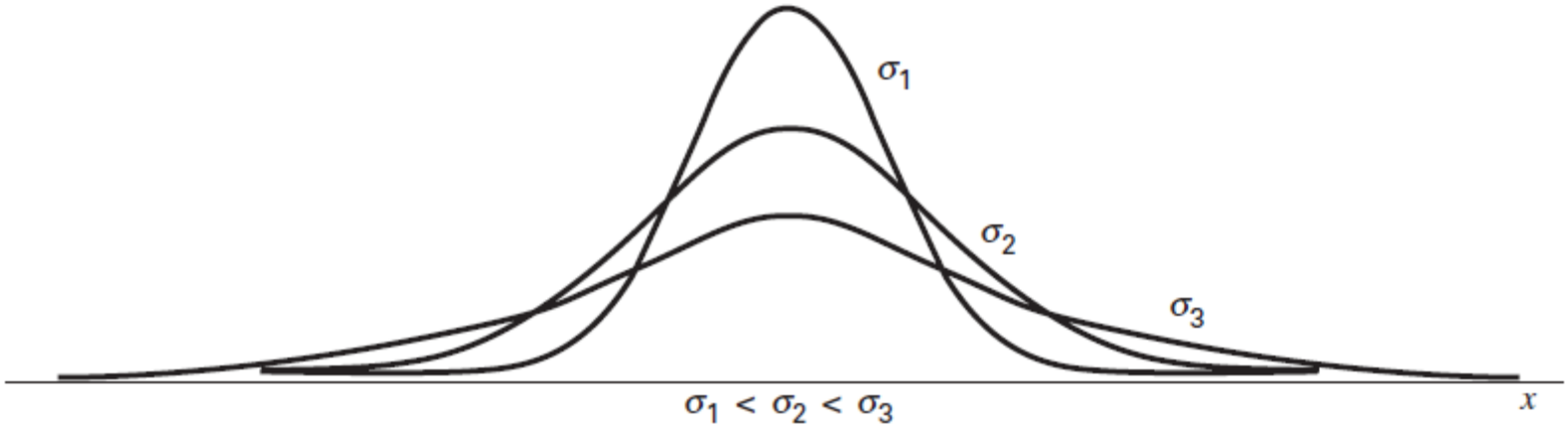


FIGURE 4.6.4 Three normal distributions with different standard deviations but the same mean.

Central Limit Theorem

- **The Central Limit Theorem** states if we draw equally sized samples from a nonnormal distribution of the means of these samples will still be normal, as long as the samples are large enough.
- How large is "large"?
- We usually say that sample size over 30 is large enough under almost all circumstances.

Example

EXAMPLE 4.7.1

The Uptimer is a custom-made lightweight battery-operated activity monitor that records the amount of time an individual spends in the upright position. In a study of children ages 8 to 15 years, Eldridge et al. (A-10) studied 529 normally developing children who each wore the Uptimer continuously for a 24-hour period that included a typical school day. The researchers found that the amount of time children spent in the upright position followed a normal distribution with a mean of 5.4 hours and standard deviation of 1.3 hours. Assume that this finding applies to all children 8 to 15 years of age. Find the probability that a child selected at random spends less than 3 hours in the upright position in a 24-hour period.

- Normal distribution with $\mu = 5.4;\ \sigma = 1.3$

- $P(X < 3) = \int_{-\infty}^{3} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ dx\ =?$

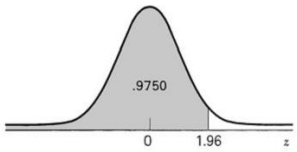
The Standard Normal Distribution

- It is a special case of normal distribution with μ=0 and σ=1
- It is obtained from normal density formula by creating a random variable z

$$z = \frac{x - \mu}{\sigma}$$

- Probabilities related to z distribution can be easily found by z-distribution tables.
- All normal distributions can be converted into the standard normal curve by defining z

TABLE D Normal Curve Areas $P(z \le z_0)$. Entries in the Body of the Table Are Areas Between $-\infty$ and z



z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.02	−0.01	0.00	z
−3.80	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	−3.80
−3.70	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	−3.70
−3.60	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002	−3.60
−3.50	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	−3.50
−3.40	.0002	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	−3.40
−3.30	.0003	.0004	.0004	.0004	.0004	.0004	.0004	.0005	.0005	.0005	−3.30
−3.20	.0005	.0005	.0005	.0006	.0006	.0006	.0006	.0006	.0007	.0007	−3.20
−3.10	.0007	.0007	.0008	.0008	.0008	.0008	.0009	.0009	.0009	.0010	−3.10
−3.00	.0010	.0010	.0011	.0011	.0011	.0012	.0012	.0013	.0013	.0013	−3.00
−2.90	.0014	.0014	.0015	.0015	.0016	.0016	.0017	.0018	.0018	.0019	−2.90
−2.80	.0019	.0020	.0021	.0021	.0022	.0023	.0023	.0024	.0025	.0026	−2.80
−2.70	.0026	.0027	.0028	.0029	.0030	.0031	.0032	.0033	.0034	.0035	−2.70
−2.60	.0036	.0037	.0038	.0039	.0040	.0041	.0043	.0044	.0045	.0047	−2.60
−2.50	.0048	.0049	.0051	.0052	.0054	.0055	.0057	.0059	.0060	.0062	−2.50
−2.40	.0064	.0066	.0068	.0069	.0071	.0073	.0075	.0078	.0080	.0082	−2.40
−2.30	.0084	.0087	.0089	.0091	.0094	.0096	.0099	.0102	.0104	.0107	−2.30
−2.20	.0110	.0113	.0116	.0119	.0122	.0125	.0129	.0132	.0136	.0139	−2.20
−2.10	.0143	.0146	.0150	.0154	.0158	.0162	.0166	.0170	.0174	.0179	−2.10
−2.00	.0183	.0188	.0192	.0197	.0202	.0207	.0212	.0217	.0222	.0228	−2.00
−1.90	.0233	.0239	.0244	.0250	.0256	.0262	.0268	.0274	.0281	.0287	−1.90
−1.80	.0294	.0301	.0307	.0314	.0322	.0329	.0336	.0344	.0351	.0359	−1.80
−1.70	.0367	.0375	.0384	.0392	.0401	.0409	.0418	.0427	.0436	.0446	−1.70
−1.60	.0455	.0465	.0475	.0485	.0495	.0505	.0516	.0526	.0537	.0548	−1.60
−1.50	.0559	.0571	.0582	.0594	.0606	.0618	.0630	.0643	.0655	.0668	−1.50
−1.40	.0681	.0694	.0708	.0721	.0735	.0749	.0764	.0778	.0793	.0808	−1.40
−1.30	.0823	.0838	.0853	.0869	.0885	.0901	.0918	.0934	.0951	.0968	−1.30
−1.20	.0985	.1003	.1020	.1038	.1056	.1075	.1093	.1112	.1131	.1151	−1.20
−1.10	.1170	.1190	.1210	.1230	.1251	.1271	.1292	.1314	.1335	.1357	−1.10
−1.00	.1379	.1401	.1423	.1446	.1469	.1492	.1515	.1539	.1562	.1587	−1.00
−0.90	.1611	.1635	.1660	.1685	.1711	.1736	.1762	.1788	.1814	.1841	−0.90
−0.80	.1867	.1894	.1922	.1949	.1977	.2005	.2033	.2061	.2090	.2119	−0.80
−0.70	.2148	.2177	.2206	.2236	.2266	.2296	.2327	.2358	.2389	.2420	−0.70
−0.60	.2451	.2483	.2514	.2546	.2578	.2611	.2643	.2676	.2709	.2743	−0.60
−0.50	.2776	.2810	.2843	.2877	.2912	.2946	.2981	.3015	.3050	.3085	−0.50
−0.40	.3121	.3156	.3192	.3228	.3264	.3300	.3336	.3372	.3409	.3446	−0.40
−0.30	.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821	−0.30
−0.20	.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207	−0.20
−0.10	.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602	−0.10
0.00	.4641	.4681	.4721	.4761	.4801	.4840	.4880	.4920	.4960	.5000	0.00

TABLE D (*continued*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	z
0.00	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359	0.00
0.10	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753	0.10
0.20	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141	0.20
0.30	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517	0.30
0.40	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879	0.40
0.50	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224	0.50
0.60	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549	0.60
0.70	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852	0.70
0.80	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133	0.80
0.90	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389	0.90
1.00	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621	1.00
1.10	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830	1.10
1.20	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015	1.20
1.30	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177	1.30
1.40	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319	1.40
1.50	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441	1.50
1.60	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545	1.60
1.70	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633	1.70
1.80	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706	1.80
1.90	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767	1.90
2.00	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817	2.00
2.10	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857	2.10
2.20	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890	2.20
2.30	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916	2.30
2.40	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936	2.40
2.50	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952	2.50
2.60	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964	2.60
2.70	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974	2.70
2.80	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981	2.80
2.90	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986	2.90
3.00	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990	3.00
3.10	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993	3.10
3.20	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995	3.20
3.30	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997	3.30
3.40	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998	3.40
3.50	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	3.50
3.60	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	3.60
3.70	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	3.70
3.80	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	3.80

Normal Distribution Applications

EXAMPLE 4.7.1

The Uptimer is a custom-made lightweight battery-operated activity monitor that records the amount of time an individual spends in the upright position. In a study of children ages 8 to 15 years, Eldridge et al. (A-10) studied 529 normally developing children who each wore the Uptimer continuously for a 24-hour period that included a typical school day. The researchers found that the amount of time children spent in the upright position followed a normal distribution with a mean of 5.4 hours and standard deviation of 1.3 hours. Assume that this finding applies to all children 8 to 15 years of age. Find the probability that a child selected at random spends less than 3 hours in the upright position in a 24-hour period.

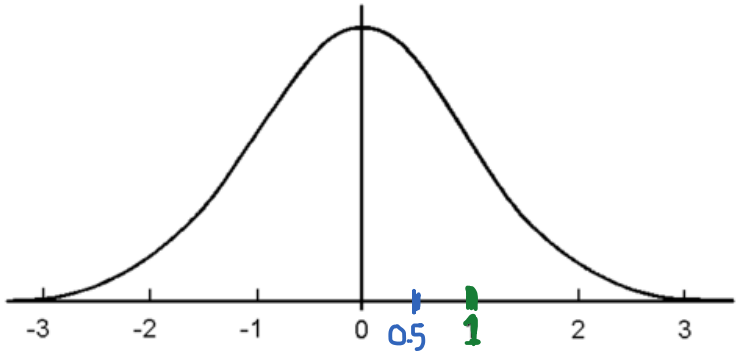
- $\mu = 5.4;\ \sigma = 1.3$

$$P(X < 3) = P\left(\frac{X - \mu}{\sigma} < \frac{3 - 5.4}{1.3}\right) = P(z < -1.85) = 0.0322$$

Example

- It is known that the results of Medical Biology and Biochemistry exams are normally distributed with $\mu_1 = 45$, $\sigma_1 = 10$ and $\mu_2 = 30$, $\sigma_2 = 2$ respectively. Ahmet had a score of 50 in Medical Biology, and 32 in Biochemistry. If the grades were assigned according to the normal curve; in which exam he did better?

- $z_1 = \frac{X_1 - \mu_1}{\sigma_1} = \frac{50-45}{10} = 0.5$
- $z_2 = \frac{X_2 - \mu_2}{\sigma_2} = \frac{32-30}{2} = 1$



- He did better in the Biochemistry exam.

The t-distribution

- It is a family of distributions based on the degrees of freedom
- It has a mean of 0
- It is symmetrical about the mean
- Variable t ranges from $-\infty$ to $+\infty$

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$

where ν is the number of *degrees of freedom* and Γ is the gamma function.

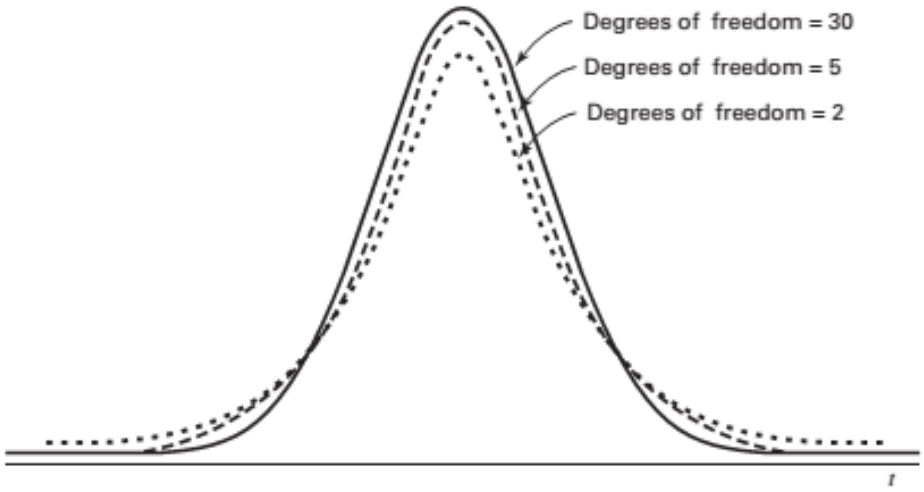


FIGURE 6.3.1 The *t* distribution for different degrees-of-freedom values.

The t-distribution

- Compared to the normal distribution, the *t* distribution is less peaked in the center and has thicker tails.

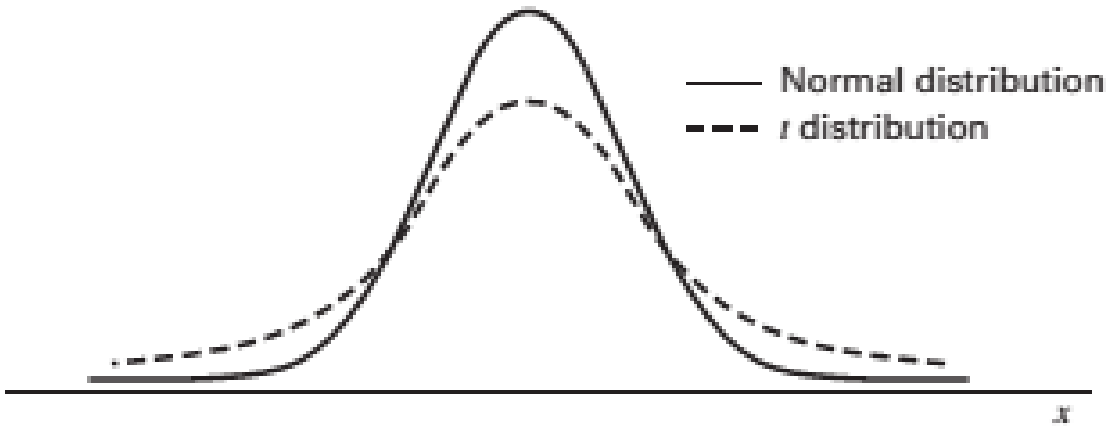
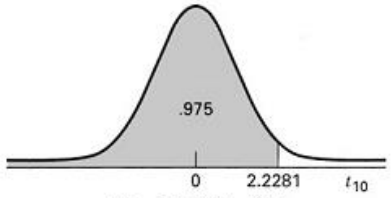


FIGURE 6.3.2 Comparison of normal distribution and *t* distribution.

TABLE E Percentiles of the t Distribution



$p(t_{10} \leq 2.2281) = .975$

d.f.	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	3.078	6.3138	12.706	31.821	63.657
2	1.886	2.9200	4.3027	6.965	9.9248
3	1.638	2.3534	3.1825	4.541	5.8409
4	1.533	2.1318	2.7764	3.747	4.6041
5	1.476	2.0150	2.5706	3.365	4.0321
6	1.440	1.9432	2.4469	3.143	3.7074
7	1.415	1.8946	2.3646	2.998	3.4995
8	1.397	1.8595	2.3060	2.896	3.3554
9	1.383	1.8331	2.2622	2.821	3.2498
10	1.372	1.8125	2.2281	2.764	3.1693
11	1.363	1.7959	2.2010	2.718	3.1058
12	1.356	1.7823	2.1788	2.681	3.0545
13	1.350	1.7709	2.1604	2.650	3.0123
14	1.345	1.7613	2.1448	2.624	2.9768
15	1.341	1.7530	2.1315	2.602	2.9467
16	1.337	1.7459	2.1199	2.583	2.9208
17	1.333	1.7396	2.1098	2.567	2.8982
18	1.330	1.7341	2.1009	2.552	2.8784
19	1.328	1.7291	2.0930	2.539	2.8609
20	1.325	1.7247	2.0860	2.528	2.8453
21	1.323	1.7207	2.0796	2.518	2.8314
22	1.321	1.7171	2.0739	2.508	2.8188
23	1.319	1.7139	2.0687	2.500	2.8073
24	1.318	1.7109	2.0639	2.492	2.7969
25	1.316	1.7081	2.0595	2.485	2.7874
26	1.315	1.7056	2.0555	2.479	2.7787
27	1.314	1.7033	2.0518	2.473	2.7707
28	1.313	1.7011	2.0484	2.467	2.7633
29	1.311	1.6991	2.0452	2.462	2.7564
30	1.310	1.6973	2.0423	2.457	2.7500
35	1.3062	1.6896	2.0301	2.438	2.7239
40	1.3031	1.6839	2.0211	2.423	2.7045
45	1.3007	1.6794	2.0141	2.412	2.6896
50	1.2987	1.6759	2.0086	2.403	2.6778
60	1.2959	1.6707	2.0003	2.390	2.6603
70	1.2938	1.6669	1.9945	2.381	2.6480
80	1.2922	1.6641	1.9901	2.374	2.6388
90	1.2910	1.6620	1.9867	2.368	2.6316
100	1.2901	1.6602	1.9840	2.364	2.6260
120	1.2887	1.6577	1.9799	2.358	2.6175
140	1.2876	1.6558	1.9771	2.353	2.6114
160	1.2869	1.6545	1.9749	2.350	2.6070
180	1.2863	1.6534	1.9733	2.347	2.6035
200	1.2858	1.6525	1.9719	2.345	2.6006
∞	1.282	1.645	1.96	2.326	2.576

Introduction to Statistical Analysis and Inference

- Confidence Intervals (interval estimates)
- Hypothesis Tests
- Correlation and Regression Analysis
- Systematic Reviews and Meta Analysis

Principles of Statistical Analysis

Principles of Statistical Analysis

- Describing the data
  - Descriptive statistics
    - Mean, standard deviation, frequency distributions
- Estimation & Confidence Intervals
- Hypothesis (significance) tests

Statistical Inference

- Estimation
- Hypothesis Testing

Statistical Inference

- Recall statistics
  - Descriptive
    - Describing the data
  - Inferential
    - Making inferences from the data
- There are two major ways of statistical inference:
  - Estimation
  - Hypothesis testing

Estimation

Point Estimate

Interval Estimate (Confidence Interval)

Estimation

- Estimation is the process of providing a value for a population parameter on the basis of information collected from a sample.
  - Point Estimate
  - Interval Estimate (Confidence Interval)

Point Estimate

- It is a single numerical value calculated to estimate the corresponding population parameter.
  - Sample mean ($\bar{X}$) is a point estimate of the population mean (μ).

- Example: The mean of the weights from a sample of patients is considered as a point estimate of the corresponding population mean.
  - $\bar{x} = 62\text{cm}$

Interval Estimate
Confidence Interval (CI)

- It is a range of values that is reasonably certain to take the value of the population parameter of interest.
- It consists of two numerical values defining a range of values that, with a specified degree of confidence, most likely includes the parameter being estimated.
- Example: The interval for the mean of the weights from a sample of patients is considered as an interval estimate of the corresponding population mean.
  - $54\text{cm} < \mu < 62\text{cm}$

Confidence Intervals

- In general, an interval estimate can be expressed as:

Estimator $\pm$ (reliability coefficient) x (standard error)

- In particular, a confidence interval for μ is:
  - $\bar{x} \pm z_{(1-\alpha/2)}\sigma_{\bar{x}}$ if σ is known
  - $\bar{x} \pm t_{(1-\alpha/2)}s_{\bar{x}}$ if σ is unknown

What do CIs tell us?

- If we construct a 95% CI for a population mean μ, it means:

Probabilistic Interpretation

- 95% of the intervals constructed with the same sample size n will include the population mean μ

Practical Interpretation

- We are 95% confident that the population mean μ is within the constructed CI.

Hypothesis Testing

Concepts of null and alternative hypotheses
One-sided or Two-sided tests of significance
Level of significance
Rejection and Acceptance Regions
Errors in Hypothesis Testing

Hypothesis Testing

- Hypothesis Testing refers to the formal procedures to reject or fail to reject statistical hypotheses.
- It is the second of two general areas of statistical inference
- In any hypothesis testing procedure, two hypotheses are considered:
  - The **null hypothesis** and The **alternative hypothesis**
- As a result of a hypothesis testing procedure, we make a decision on the null hypothesis.
  - Reject or Fail to reject

Null Hypothesis

- Null hypothesis
  - denoted by H_0,
  - is an assumption about a relevant population parameter. (e.g, mean, proportion, etc. of a population).
  - It is not about a sample, and sample statistic is not used in formulating the null hypothesis.
- Null hypothesis is usually a hypothesis of ***no difference***.
- This is the main hypothesis which we wish to test, since acceptance of it commonly implies “no difference”, “no effect”, “no relation”, “independence (no dependency)” etc.

Alternative Hypothesis

- Alternative hypothesis
  - denoted by H_1 or H_A
  - is a claim that the null hypothesis is false, so the population parameter takes on a value different from the value or values specified by the null hypothesis.

- We do not actually make a decision on the alternative hypothesis, we make decision on the null hypothesis.

Rules for stating statistical hypotheses

- If we test for the difference, the null hypothesis should contain a statement of equality,
  - either $=$, $\leq$, or $\geq$
- The null hypothesis is the hypothesis that is tested
- The null and alternative hypotheses are complementary.

Example

- Can we conclude that a certain population mean is not 50?
  - $H_0: \mu = 50$
  - $H_1: \mu \neq 50$

  OR

  - H_0 : the population mean equals 50
  - H_1: the population mean is not equal to 50

Example

- Can we conclude that a certain population mean is not 50?
  - $H_0: \mu = 50$
  - $H_1: \mu \neq 50$

This is a **two-sided hypothesis test**
If μ is not equal to 50, it might be <50 or >50

OR

- H_0 : the population mean equals 50
- H_1: the population mean is not equal to 50

Example

- We want to know if we can conclude that the population mean is greater than 50.
  - $H_0: \mu \leq 50$
  - $H_1: \mu > 50$

  OR

  - H_0 : the population mean is less than or equal to 50
  - H_1: the population mean is greater than 50

Example

- We want to know if we can conclude that the population mean is greater than 50.
  - $H_0: \mu \leq 50$
  - $H_1: \mu > 50$

This is a **one-sided hypothesis test**

OR

  - H_0 : the population mean is less than or equal to 50
  - H_1: the population mean is greater than 50

Example

- Suppose data is collected from a sample about heart disease and diabetes; and we want to test if these two diseases are independent of each other or not.
  - Is it possible to state hypotheses in mathematical notations?
    - NO!
  - So, we need to state hypotheses in words only:

H_0 : heart disease and diabetes are independent of each other

H_1: heart disease and diabetes are dependent to each other

This is a **two-sided hypothesis test**
Since It has no direction

Terms in Hypothesis Testing

- Test statistic
- Significance Level
- Types of Errors
- p-values

Test Statistic

- $test\ statistic = \frac{relevant\ statistic - hypothesized\ parameter}{standard\ error\ of\ the\ relevant\ statistic}$
- Example: A sample with size $n = 124$ was taken from a population and their hemoglobin levels were recorded. Sample mean and sample standard deviation were calculated as $\bar{x} = 14$; $s = 5.7$
- If we conduct a t-test in order to determine if the hemoglobin levels of the corresponding population is grater than 12.5;
- The test statistic t is calculated by:
  - $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{14 - 12.5}{5.7/\sqrt{124}} = 2.93$

Distribution of the test statistic

- Based on the distribution of the test statistic, the corresponding table value for the test statistic is found.

Significance Level

- When we try to make inferences about a population by using sample data, It is not possible to prove or disprove either H_0 or H_1.

- Rather, it is possible to make a decision based on probability by accepting making an incorrect decision with a small probability.

- The decision about to reject/fail to reject H_0 is made on the basis of the desired ***level of significance*** (α)

Significance Level

- Defn: The level of significance α is the probability of rejecting a true null hypothesis
  - It is the probability of wrongly rejecting H_0
- We select a small value of α in order to make the probability of rejecting a true null hypothesis small.
- Most frequently encountered values for α are: 0.01, 0.05, 0.1
- Choosing $\alpha = 0.05$ means that, we accept to be wrong 5 times out of 100 tests.

Significance Level

- Why mostly $\alpha = 0.05$ is used?
  - In 1925; Ronald Fisher, a famous statistician and geneticist, proposed the level p=0.05 by considering 1 in 20 is a reasonable standard for rarity.

Types of Errors

- There are two possible types of error in hypothesis testing:
  - Type I error and Type II error

- Type I error, α, is committed when a true null hypothesis is rejected.
  - Also called as significance level

- Type II error, β, is committed when a false null hypothesis is not rejected.

Types of Errors

- We are able to control type I error by choosing α small. However, generally we have no control over β.

- Since the true state of affairs is unknown, we never know whether we have committed one of these errors when we reject or fail to reject a null hypothesis

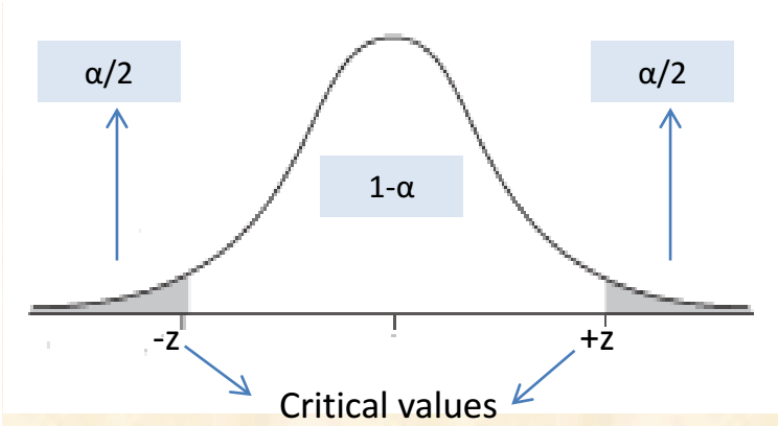
		Condition of Null Hypothesis	
		True	False
Possible Action	Fail to reject H_0	Correct action	Type II error β
	Reject H_0	Type I error α	Correct action

FIGURE 7.1.1 Conditions under which type I and type II errors may be committed.

Decision Criteria for Hypothesis tests based on rejection-nonrejection regions

- For two-sided hypotheses:

Rule: the area of the rejection region $= \alpha$

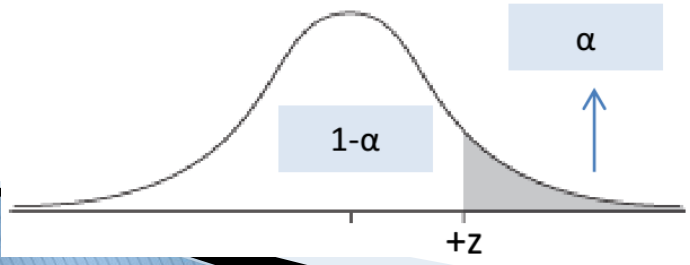


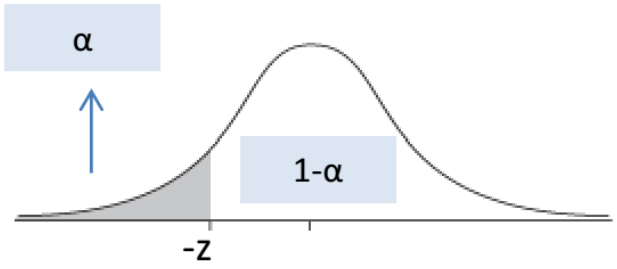
If calculated test statistic falls into the rejection region, we reject H_0
Otherwise, we fail to reject H_0

Rejection region (or) critical region

Acceptance region

- For one-sided hypotheses:





p-value

- Defn: A p-value is the probability that the computed value of a test statistic is at least as extreme as a specified value of the test statistic when H_0 is true.

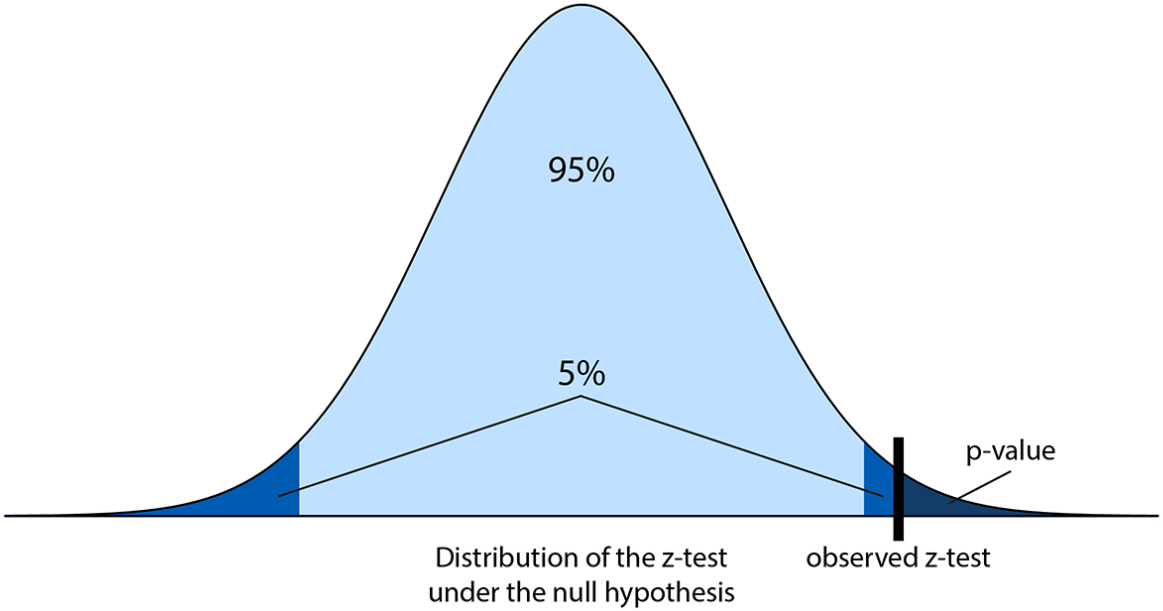
- In other words, the p-value is the smallest value of α for which we can reject a null hypothesis.

- p-value measures the strength of evidence in support of H_0
  - Example: For α=0.05, we reject H_0 more strongly if p=0.002 rather than getting p=0.042

Decision Criteria for Hypothesis tests based on p-value

- If the p-value is less than the significance level, then we reject the null hypothesis H_0
  - Reject H_0 if $p < \alpha$

- The smaller the p-value, the stronger the evidence against H_0



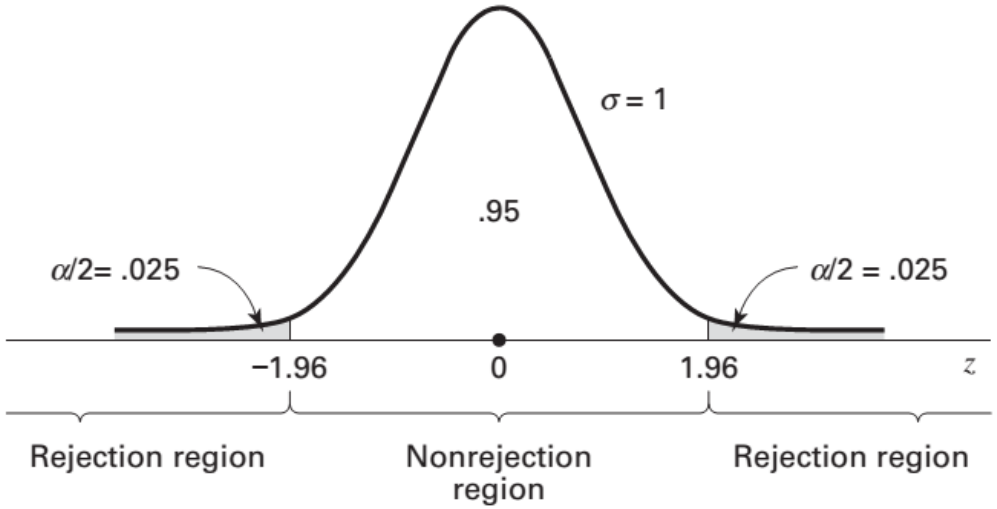
Steps in Hypothesis Testing

- 1. Evaluate data
- 2. Review assumptions
  - parameters related to the population, distribution
- 3. State hypotheses
- 4A. Select and calculate test statistics
- 4B. Find p-value
- 5. State decision rule
- 6. Make statistical decision:
  - Reject H_0 (conclude H_1 is true)
  - Fail to reject H_0 (conclude H_0 might be true)
- 7. Conclusion

Example

- Researchers are interested in the mean age of a certain population
- A random sample of 10 individuals drawn from the population of interest has a mean of 27.
- Assuming that the population is approximately normally distributed with a variance of 20, can we conclude that the mean is different from 30 years? (Let $\alpha = 0.05$)

- 1) Evaluate data
  - $n = 10$, $\bar{x} = 27$
- 2) Assumptions
  - Population approximately normally distributed
  - $\sigma^2 = 20$, $\alpha = 0.05$
- 3) Hypotheses: (two-sided)
  - $H_0: \ \mu = 30$
  - $H_1: \ \mu \neq 30$
- 4A) Test Statistic:
  - $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{27-30}{\sqrt{20}/\sqrt{10}} = -2.12$
- 4B) p-value:
  - $p = 0.0340$



- 5) Decision Rule:
  - Reject H_0 if $|z| < 1.96$
- 6) Statistical Decision
  - by test statistic
    - $-2.12 < -1.96$ -> Reject H_0
  - by p-value
    - $p = 0.0340 < 0.05$ -> Reject H_0
- 7) Conclusion:
  - With 95% significance, we conclude that the population mean is not equal to 30

Some practice in SPSS

- Descriptive Statistics
  - Frequency analysis
  - Finding descriptive measures
  - Creating crosstabs
  - Finding descriptive measures for grouped data
  - Transform a variable

How to determine the appropriate statistical method?

- Statistical Inference
  - Estimation
  - Hypothesis Tests
  - Risk Comparison
    - Relative Risk / Odds Ratio
  - Correlation
  - Regression
  - Meta Analysis

How to determine which hypothesis test to use?

- When deciding on the appropriate hypothesis test, we should consider:

1) The type of the data
2) The distribution of the data
3) The number of groups of observations
4) If there are groups, the independence/dependence of the group of observations

			Parametric Methods **Assumption: Data have to be sampled from Gaussian distribution (normally distributed)&populations have equal variances (homogenity of variances)**	**Non parametric Methods** **No assumption. It is preferred when normallity is violated and having small sample size (n<30)**
Quantitative Data	**1 sample vs hypothesized value comparison**		**One-sample t-test**	Signed rank test or Wilcoxon Rank Sum Test
	2 samples Comparisons	*Independent Samples*	**Unpaired t-test**	**Mann-Whitney U Test**
		Paired Samples or Matched Pair Samples	**Paired t-test**	**Wilcoxon Matched Pairs Test**
	More than 2 samples Comparisons	*Independent Samples*	**One way-ANOVA Test** If p<0.05 according to F-test, Multiple Comparison Tests are needed; such as; Tukey, Bonferroni, Scheffé Multiple Comparison Tests To compare all vs. only with control group, Dunnett Multiple Comparison Test can be used.	**Kruskal-Wallis Test** If p<0.05 use Dunn Multiple Comparison Test
		Paired Samples or Matched Pair Samples	**Repeated Measures ANOVA Test** If p<0.05 according to F-test, Multiple Comparison Tests are neededStudent such as Newman-Keuls	**Friedman Test** If p<0.05 use Dunn Multiple Comparison Test
Categorical (qualitative) Data	**Crosstabulation**	*Independent Samples*	---	**X^2-Test * Homogenity test** ***Independence test** If the association is significant ; Phi or Cramer's V coefficients can be used İn order to measure the degree and to determine the direction of the association.
		Paired Samples or Matched Pair Samples	---	**McNemar X^2 test (2x2)** or **Stuart-Maxwell testi(3x3)**
Correlation Coefficient (This *measures the degree* and *determines the direction* of linear co-relation between two coninous variables, x and y)			**Pearson Correlation Coefficient**	**Spearman Correlation Coefficient**

Parametric vs. Nonparametric methods for quantitative data

- Parametric tests
  - involve estimating population parameters such as the mean
  - are based on the assumptions of normality or known variances (stringent assumptions)
- What if data does not follow a Normal distribution?
  - Then use a nonparametric test
- Nonparametric tests
  - were developed for these situations where no (or fewer) assumptions have to be made
  - are also called as Distribution-free tests
  - Still have assumptions but they are less stringent
  - Can be applied for a normally distributed data, but Parametric tests have greater power IF the assumptions met

Parametric Methods for Quantitative data

- Provide us inference about mean(s)
- Assumptions:
  - Data have to be ratio or interval scaled
  - Data has to be approximately normally distributed
    - For two independent groups, data in both groups should be normally distributed
    - For paired data, the difference between pairs should have normal distribution
  - Variances have to be homogeneous
    - For two or more independent groups, the variance should be homogeneous among the groups (this can be checked by F-test)

Parametric Tests for Quantitative data

			Parametric Methods **Assumption: Data have to be sampled from Gaussian distribution (normally distributed)&populations have equal variances (homogenity of variances)**
Quantitative Data	**1 sample vs hypothesized value comparison**		**One-sample t-test**
	2 samples Comparisons	*Independent Samples*	**Unpaired t-test**
		Paired Samples or Matched Pair Samples	**Paired t-test**
	More than 2 samples Comparisons	*Independent Samples*	**One way-ANOVA Test** If $p<0.05$ according to F-test, Multiple Comparison Tests are needed; such as; Tukey, Bonferroni, Scheffé Multiple Comparison Tests To compare all vs. only with control group, Dunnett Multiple Comparison Test can be used.
		Paired Samples or Matched Pair Samples	**Repeated Measures ANOVA Test** If $p<0.05$ according to F-test, Multiple Comparison Tests are neededStudent such as Newman-Keuls

t-tests

1) One sample t-test
   - We are interested in whether the mean of the variable in the population of interest differ from a specific hypothesized mean.
   - Comparison parameter has been estimated from prior research or is derived from theory.

2) Independent samples (or Unpaired) t-test
   - We compare the means of two independent populations

t-tests

- 3) Paired samples t-test
  - We compare two dependent groups.
  - Paired data may arise
    - when the same individuals are studied more than once, usually in different circumstances.
    - when we have two different groups of subjects who have been individually matched.
  - In typical paired studies, first measurements are taken to establish a baseline (**before** measurement); then, after some intervention or time, the same subjects are measured again (**after** measurement).
  - In a paired samples t-test, we are interested in the difference between the observations for each individual.

Example: One Sample t-test

- A researcher is planning a psychological intervention study, but before he proceeds he wants to characterize his participants' depression levels. He tests each participant on a particular depression index, where anyone who achieves a score of 4.0 is deemed to have 'normal' levels of depression. Lower scores indicate less depression and higher scores indicate greater depression. He has recruited 40 participants to take part in the study. Depression scores are recorded in the variable dep_score. He wants to know whether his sample is representative of the normal population (i.e., do they score statistically significantly differently from 4.0).

- Data: 3.68, 3.98, 3.72, 3.98, … , 4.26, 3.51, 4.04 (n=40)

Example: One Sample t-test

- The example is a one sample design to compare the population mean with a known threshold.

- Can we directly apply one-sample t-test?
  - No!
  - First, we need to check if the assumption of normality is met.

Normality check

- In order to check if the data is normally distributed or not, Shapiro-Wilk test is conducted.
  - H_0: Data is normally distributed
  - H_1: Data is not normally distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dep_score	.103	40	.200*	.958	40	.138

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

P-value$=0.138 > 0.05$

We fail to reject H_0

Data is normally distributed

Example: One Sample t-test

- Since data is normally distributed, one sample t-test can be conducted:
  - H_0: The mean depression score of the study population is not different than 4 ($\mu_{dep\_score} = 4$)
  - H_1: The mean depression score of the study population is different than 4 ($\mu_{dep\_score} \neq 4$)

Example: One Sample t-test

- Results:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
dep_score	40	3.5158	.58434	.09239

One-Sample Test

	Test Value = 4					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
dep_score	-5.241	39	.000	-.48425	-.6711	-.2974

P-value=0.000 $<$ 0.05

We reject H_0

$\mu_{dep\_score} \neq 4$, it is less than 4

Lower by 0.48

CI of the difference is (-0.67,-0.3)

Example: One Sample t-test

- How to report these results?

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
dep_score	40	3.5158	.58434	.09239

One-Sample Test

	Test Value = 4					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
dep_score	-5.241	39	.000	-.48425	-.6711	-.2974

- Mean depression score (3.52 ± 0.58) was lower than the normal depression score of 4.0, with a statistically significant difference of 0.48 (95% CI, 0.3 to 0.67), $t(39) = -5.241$, $p < 0.0005$.

Example: Independent Samples t-test

- The table below shows the cholesterol levels of 20 patients from City A, and 18 patients from City B. Are the means of cholesterol levels from these two cities statistically different from each other?

	cholesterol levels (mmol/l)								
	1	2	3		17	18	19	20	
city A	4.1	5.3	4.9	…	5.2	6.4	5.1	5.8	(n=20)
city B	6.6	5.2	5.9	…	6.3	6.4			(n=18)

Example: Independent Samples t-test

- The example is a two independent samples design to compare the population means
- Can we directly apply independent samples t-test?
  - No!
  - First, we need to check the assumptions
    - Observations in each group should follow a normal distribution
      - Shapiro-Wilk test
    - The variances in the two samples should be equal(homogeneous)
      - Levene's test for homogeneity of variances (this test is done within the procedure of t-test / so no extra work needed)

Normality check

- In order to check if the data in groups is normally distributed or not, Shapiro-Wilk test is conducted.
  - H_0: Data is normally distributed
  - H_1: Data is not normally distributed

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	group	Statistic	df	Sig.	Statistic	df	Sig.
Cholesterol (mmol/l)	City A	.117	20	.200*	.973	20	.815
	City B	.165	18	.200*	.932	18	.215

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

For City A: p-value=0.815 $>$ 0.05

We fail to reject H_0 ; data from City A is normal

For City B: p-value=0.215 $>$ 0.05

We fail to reject H_0 ; data from City B is normal

Example: Independent Samples t-test

- Since data is normally distributed, independent samples t-test can be conducted:
  - H_0: The mean cholesterol levels in City A is not different than the mean cholesterol level in City B $(\mu_{City\ A} = \mu_{City\ B})$
  - H_1: The mean cholesterol levels in City A is different than the mean cholesterol level in City B $(\mu_{City\ A} \neq \mu_{City\ B})$

Example: Independent Samples t-test

- Results:

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Cholesterol (mmol/l)	City A	20	5.670	.7740	.1731
	City B	18	6.006	.6958	.1640

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Cholesterol (mmol/l)	Equal variances assumed	.522	.475	-1.399	36	.170	-.3356	.2398	-.8219	.1508
	Equal variances not assumed			-1.407	36.000	.168	-.3356	.2384	-.8191	.1480

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Cholesterol (mmol/l)	Equal variances assumed	.522	.475	-1.399	36	.170	-.3356	.2398	-.8219	.1508
	Equal variances not assumed			-1.407	36.000	.168	-.3356	.2384	-.8191	.1480

$H_0: {\sigma_1}^2 = {\sigma_2}^2$
$H_1: {\sigma_1}^2 \neq {\sigma_2}^2$

$p = 0.475 > 0.05$

Fail to reject H_0
Variances are equal

$H_0: \mu_{City\ A} = \mu_{City\ B}$
$H_1: \mu_{City\ A} \neq \mu_{City\ B}$

$p = 0.17 > 0.05$

Fail to reject H_0
Means are equal

Mean cholesterol level in City A(5.67 ± 0.77) does not significantly differ from the mean cholesterol level in City B (6 ± 0.7), $t(36) = -1.399$, $p = 0.17$.

Example: Paired Samples t-test

- In a tumor size study, two doctors were shown the same set of tumor pictures. The volume of tumor was measured (in cm^3) by two separate physicians under similar conditions, and the following values are obtained. Is there a difference in tumor volume measurement based on physician?

Tumor	Dr1	Dr2
1	15.8	17.2
2	22.3	20.3
3	14.5	14.2
4	15.7	18.5
5	26.8	28.0
6	24.0	24.8
7	21.8	20.3
8	23.0	25.4
9	29.3	27.5
10	20.5	19.7

Example: Paired Samples t-test

- The example is a paired samples design to compare the population means

- Can we directly apply paired samples t-test?
  - No!
  - First, we need to check the assumption of normality:
    - The differences should be normally distributed
      - Shapiro-Wilk test

Normality check

- In order to check if the data is normally distributed or not, Shapiro-Wilk test is conducted.
  - H_0: Data is normally distributed
  - H_1: Data is not normally distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
difference	.138	10	.200*	.930	10	.452

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

P-value$=0.452 > 0.05$

We fail to reject H_0

Data is normally distributed

Example: Paired Samples t-test

- Since data is normally distributed, paired samples t-test can be conducted:
  - H_0: There is no difference in tumor volume measurement based on physician $(\mu_{dr1} = \mu_{dr2})$
  - H_1: There is a difference in tumor volume measurement based on physician $(\mu_{dr1} \neq \mu_{dr2})$

Example: Paired Samples t-test

- Results:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	dr1	21.3700	10	4.87762	1.54244
	dr2	21.5900	10	4.60880	1.45743

High positive correlation ($r = 0.934, p < 0.0005$),

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	dr1 & dr2	10	.934	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	dr1 - dr2	-.22000	1.74407	.55152	-1.46763	1.02763	-.399	9	.699

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	dr1 - dr2	-.22000	1.74407	.55152	-1.46763	1.02763	-.399	9	.699

$H_0: \mu_{dr1} = \mu_{dr2}$
$H_1: \mu_{dr1} \neq \mu_{dr2}$

$p = 0.699 > 0.05$

Fail to reject H_0
Means are equal

We can conclude that mean volume measurement by physician1 $21.37 \pm 4.88)$ is not significantly different than the mean volume measurement by physician2 $21.59 \pm 4.61)$; $t(9) = -1.47$; $p = 0.699$