Review of Probability and Statistics

Outline

- **Probability**
  - Random variables and probability
  - Discrete distributions
  - Continuous distributions
  - Joint probability of multiple random variables
- **Statistics**
  - Sampling
  - Statistical inference
  - Estimation
  - Hypothesis testing

Sample Space and Event

- We run an experiment whose outcome is uncertain
  - e.g., Toss a coin

- **Sample space (Ω):** the set of all possible outcomes of an experiment
  - Experiment 1 - Toss a coin: $\Omega = \{H, T\}$
  - Experiment 2 - Toss a coin twice: $\Omega = \{HH, HT, TH, TT\}$

- **Event (E):** any collection (subset) of the outcomes of sample space
  - Experiment 2 (Toss a coin twice)
    - The 1st toss H: $E = \{HH, HT\}$
    - No tail: $E = \{HH\}$
    - At least one H: $E = \{HH, HT, TH\}$

Set Theory

- **Complement** of an event A, (A'): the set of all outcomes in the sample space Ω, that are not contained in A

- **Union** of A and B ($A \cup B$): the event consisting of all outcomes that are either in A or in B or in both events

- **Intersection** of A and B ($A \cap B$): the event consisting of all outcomes that are in *both* A and B.

- **Mutually exclusive**
  - ∅ denote the null event
  - $A \cap B = \emptyset$

Probability Axioms

- For any event A, $P(A) \geq 0$
- $P(\Omega) = 1$
- If $A_1, A_2, A_3, \ldots$ is an infinite collection of disjoint events, then

$$P\left(A_1 \cup A_2 \cup A_3 \ \ \ldots\right) = \sum_{i=1}^{\infty} P(A_i)$$

- $P(\emptyset) = 0$

Probability Properties

- $P(A) + P(A') = 1$
- $P(A) \leq 1$
- For any two events A and B

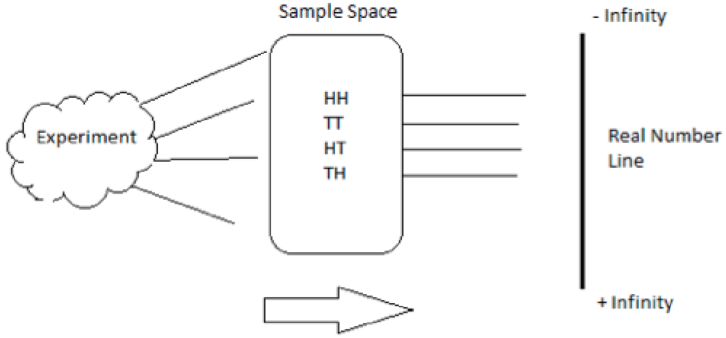
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

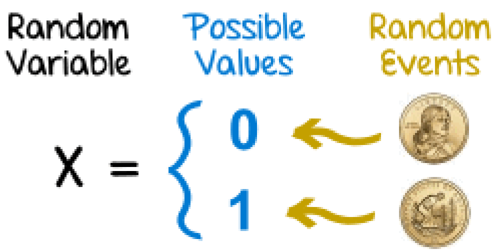
- For any three events A, B and C

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap \mathrm{C})$$

Random Variable (RV)

- A function or a code that maps simple events to a real number
- $X: \Omega \rightarrow \mathrm{R}$

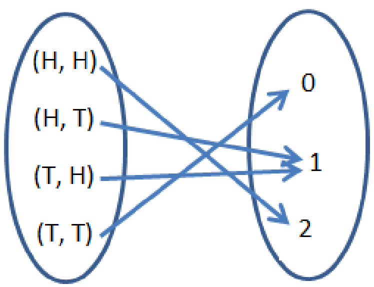




http://www.mathsisfun.com/data/random-variables.html

- Many ways to code!

Toss 2 coins



(Biobio.loc.edu)

Sample Space (S)	Random Variable (X)
HH	0
HT	1
TH	2
TT	3

(www.rfortraders.com)

Example: Toss 3 Coins

- **Why study this?**
  - To get the probabilities of various events of interest
  - Assess risk and your bet
- **Let us code it in the following way:**
  - X = "The number of Heads"
  - $X(S) = \{0, 1, 2, 3\}$
- **Any problem here?**
  - Good for counting: Probability of heads
    - $\mathrm{P}(X=0) = 1/8$; $\mathrm{P}(X=2) = 3/8$;
    - $\mathrm{P}(X>1) = 1/2$; $\mathrm{P}(X>\text{ or }<2) = 5/8$
  - Not so for the order:
    - Probability that the first toss is a head

Types of Random Variables

- **Discrete**
  - Integer coding (take finite or countable number of values)
  - X maps to the integer line
  - E.g., number of people waiting in the post office
- **Continuous**
  - Real number coding
  - X maps to real line
  - E.g., height of students in the class
- **Univariate vs. Multivariate RV**
  - Scalar vs. vector coding
  - Two tosses of a coin-
    - Univariate RV: $X =$ # of heads
    - Multivariate (here, bivariate) RV: X=["Is 1st toss H?", "Is 2nd toss H?"]

Discrete Distributions

- **Discrete probability distribution**
  - Defined on discrete rv
  - Probability mass function (**PMF**)

- **Typical distributions**
  - Discrete uniform
  - Bernoulli
  - Binomial
  - Geometric
  - Poisson
  - Negative binomial
  - Hyper-geometric

Discrete Uniform Distribution

- **Experiment**
  - One trial
  - *k* possible outcomes
  - All outcomes equally probable
- **Random variable**: X – outcome of the trial
- **Probability distribution**:

$$p(x) = \begin{cases} \frac{1}{k} & x \in S \\ 0 & \text{otherwise} \end{cases}$$

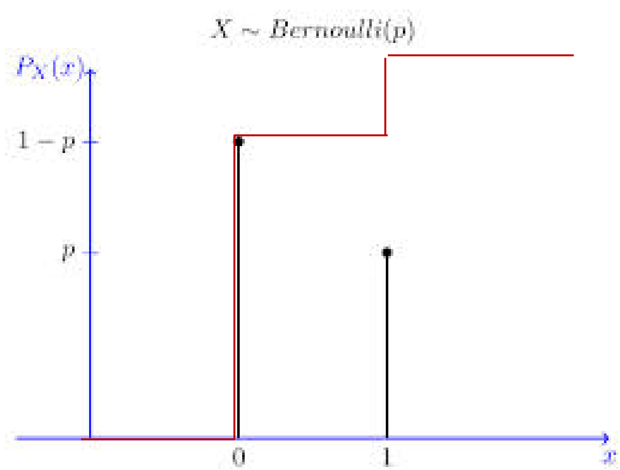
- **Example**: toss a fair die ($k = 6$)
  - S = {1, 2, 3, …,6}
  - Expectation: $E(X) = \frac{1}{k}\sum_{i=1}^{k} x_i$
  - Variance: $V(X) = \frac{1}{k}\sum_{i=1}^{k}(x_i - E(X))^2$

Bernoulli Distribution

- **Experiment**
  - A single ($n = 1$) trial with two possible outcomes ("success" and "failure")
  - $P(\{\text{success}\}) = p$
- **Random variable**: X – outcome of the trial (1 or 0)
- **Probability distribution**
  - **Probability mass function (PMF):** $P(X = x) = p(x)$ $\quad p(1) = p,\ p(0) = 1 - p$
  - **Cumulative distribution function (CDF)**

$$F(x) = P(X \le x) = \sum_{z=-\infty}^{x} p(z)$$

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - p & x \in [0,1) \\ 1 & x \ge 1 \end{cases}$$



- **Expectation**

$$\mu_X = E(X) = \sum_{x=-\infty}^{\infty} x p(x) = p$$

- **Variance**

$$\sigma_X^2 = V(X) = E[(X - \mu_X)^2] = p(1 - p)$$

- **Example**: toss a fair coin once

Binomial Distribution

- **Experiment**
  - n repeated **independent** trials
  - Each trial has two possible outcomes ("success" and "failure")
  - $P\left(\{i^{th} \text{ trial is success}\}\right) = p$ for all i
- **Random variable:** X – number of successful trials
- **PMF** $P(X = x)$

$$b(x; n, p) = \binom{n}{x} p^x (1 - p)^{n-x}$$

- **CDF** $P(X \leq x)$

$$B(x; n, p) = \sum_{r=0}^{x} \binom{n}{r} p^r (1 - p)^{n-r}$$

- **Expectation:** $E(X) = np$
- **Variance:** $\mathrm{var}(X) = np(1 - p)$

Geometric Distribution

- **Experiment**
  - Indeterminate number of repeated trials
  - Each trial has two possible outcomes ("success" and "failure")
  - $P(\{the\ outcome\ of\ the\ i^{th}\ \text{trial is success}\}) = p$ for all i
  - **Independent** trials
- **Random variable:** X – number of trials until 1st success
- **Probability distribution** (PMF): $P(x) = p(1-p)^{x-1}$
- **Expectation & variance**

$$E(X) = \frac{1}{p} \qquad \text{var}(X) = \frac{1-p}{p^2}$$

- **Example:** repeated attempts to start an engine; play a lottery until you win

Negative Binomial Distribution

- **Experiment**
  - Indeterminate number of repeated trials
  - Each trial has two possible outcomes (success and failure)
  - $P\left(\{the\ outcome\ of\ the\ i^{th}\ \text{trial is success}\}\right) = p$ for all i
  - Independent trials
  - Keep going until the r^{th} success
- **Random variable:** X – #trials until r successes
- **Probability distribution** (PMF)

$$b^*(x; r, p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

- **Expectation and variance**

$$E(X) = \frac{r}{p} \qquad \text{var}(X) = \frac{r(1-p)}{p^2}$$

- **Example:** fabricating r defective computer chips

Hyper-geometric Distribution

- **Experiment:**
  - A random sample of size n is selected from N items
  - There are k items of one type (success) and $N - k$ items of another type (failure)
- **Random variable:** X – number of success selected
- **Probability distribution** (PMF)

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

- **Expectation & variance**

$$E(X) = \frac{nK}{N} \qquad \mathrm{var}(X) = \frac{N-n}{N-1}\frac{nK}{N}\left(1 - \frac{k}{N}\right)$$

- **Example:** select a random sample of 5 spark plugs from a batch of 40 of which 3 are defective

Poisson Distribution

- **Experiment:** recurring trials in space or time
  - The events occur at a point in time or space
  - The number of events occurring in one region is **independent** of the number occurring in any disjoint region
  - Probability of n events in region/interval 1 **=** Probability of n events in region/interval 2, when the two regions/intervals have the same size
- **Random variable:** number of events occurring in the given time interval or region of space
- **Probability distribution** (PMF)

$$\text{Poisson}(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!}$$ (λ: average number of events in the region/interval)

- **Expectation & variance:** $E(X)=\lambda$ $\quad$ $\text{var}(X)=\lambda$
- **Example:** number of emails arriving in a specified (1 hour) period; number of arrived jobs

Continuous Distributions

- **Continuous probability distribution**
  - Defined on continuous rv
  - Probability density function (**PDF**)

- **Typical distributions**
  - Continuous uniform
  - Exponential
  - Gamma
  - Normal

Continuous Uniform Distribution

- **Definition:** A continuous RV X is said to have a uniform distribution on the interval $[a, b]$ if the PDF of X is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

- **Expectation & variance**

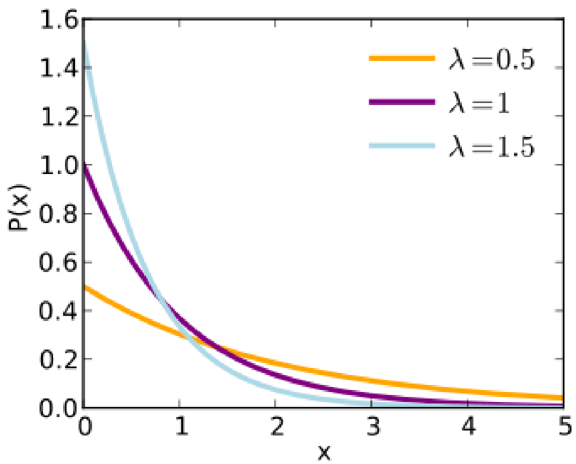
$$E(X) = \frac{a+b}{2} \qquad \mathrm{var}(X) = \frac{(b-a)^2}{12}$$

- **Example:** Spin the dial so that it comes to rest at a random position. Find the probability that the dial will land somewhere between 5 and 300.

Exponential Distribution

- **Definition:** Let λ be a positive real number, RV X is called an exponential RV ($X \sim \exp(\lambda)$) if

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$



- **Expectation & variance**

$$E(X) = \frac{1}{\lambda} \qquad \mathrm{var}(X) = \frac{1}{\lambda^2}$$

- **Example:** often used to model life time of products, waiting time, time between random events.

Gamma Distribution

- **Definition:** A continuous RV X is said to have a gamma distribution ($X \sim \text{gamma}(\alpha, \beta)$, $\alpha > 0, \beta > 0$) if

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- **Exponential distribution:** $\alpha = 1$ and $\beta = \frac{1}{\lambda}$
- **Expectation & variance**

$$E(X) = \alpha\beta \qquad \text{var}(X) = \alpha\beta^2$$

- **Example:** time until event occurs for α times

Normal (Gaussian) Distribution

- **Definition:** A continuous RV X is said to have a normal distribution ($X \sim N(\mu, \sigma^2)$) with parameter μ ($-\infty < \mu < \infty$) and σ ($\sigma > 0$), if

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

- **Standard normal distribution/RV Z:** $\mu = 0$ and $\sigma = 1$

$$f(z; 0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} (-\infty < z < \infty)$$

- **CDF of Z:** $P(Z \le z) = \int_{-\infty}^{z} f(y; 0,1) dy$ (often denoted by $\Phi(z)$)

Joint Probability Mass Function

- If X and Y are two discrete rv's defined on $\mathcal{S}$, the sample space for an experiment, their joint probability mass function is

$$p(x, y) = P(X = x \text{ and } Y = y)$$

- The marginal probability mass functions of X and Y are

$$p_x(x) = \sum_y p(x, y) \quad \text{and} \quad p_y(y) = \sum_x p(x, y)$$

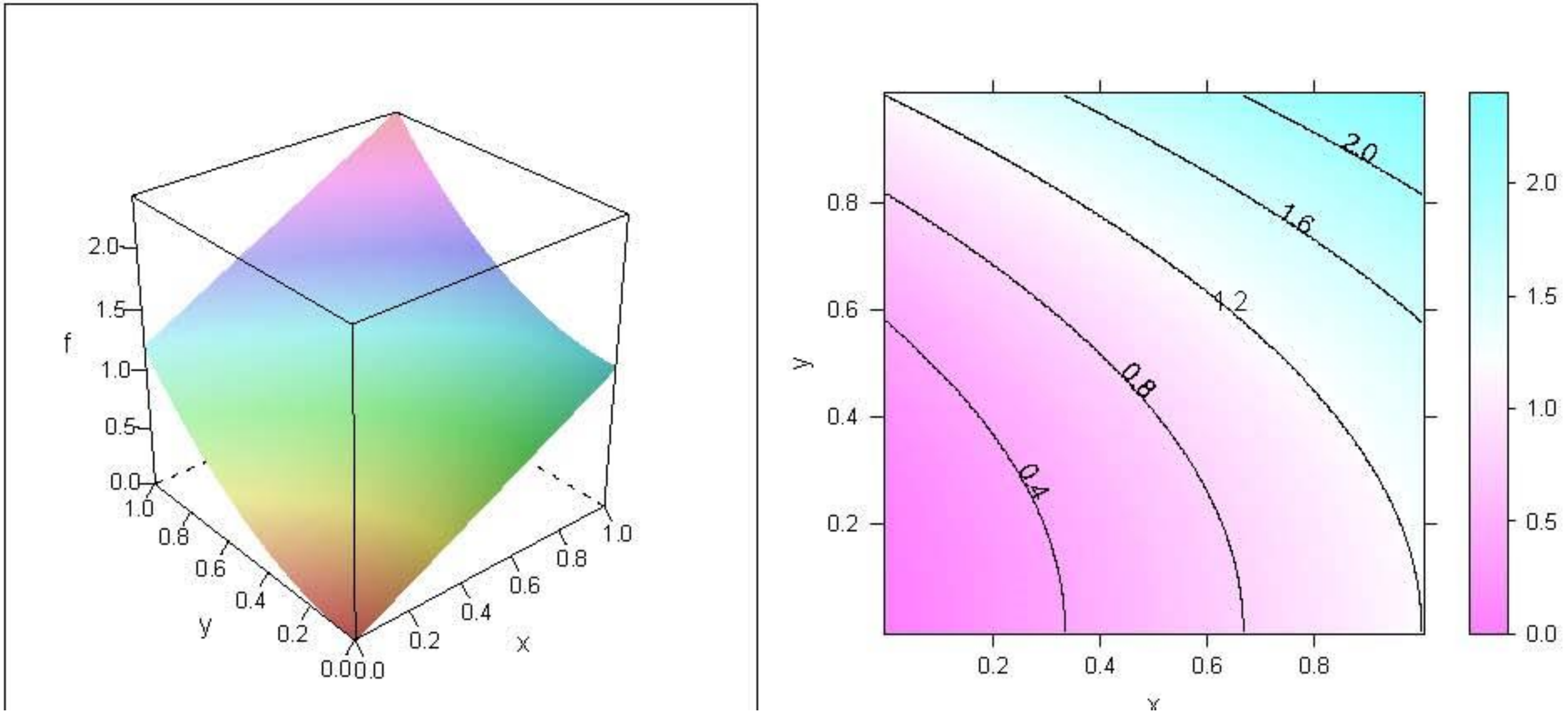
Joint Probability Density Function

- If X and Y are two continuous rv's then $f(x.y)$ is their joint density function if
$$P[(X,Y) \in A] = \int \int_A f(x,y)\,dx\,dy$$
- The marginal probability density functions of X and Y are
$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad \text{and} \quad f_y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

Example of joint probability density

Example 5.3 describes a joint probability distribution with density

$$f(x,y) = \begin{cases} \frac{6}{5}\left(x+y^2\right) & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$



Conditional distributions

- For continuous random variables X and Y with joint pdf $f(x,y)$ and marginal pdfs $f_X(x)$ and $f_Y(y)$, the conditional probability density of Y, given $X = x$ is

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} \qquad -\infty \leq y \leq \infty$$

  provided that $f_X(x) > 0$.
- For discrete random variables X and Y with joint pmf $p(x,y)$ and marginal pmfs $p_X(x)$ and $p_Y(y)$ the conditional pmf of Y given $X = x$ is

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$$

  provided that $p_X(x) > 0$.

Independent Random Variables

- Discrete random variables X and Y are said to be independent if

$$p(x, y) = p_X(x) \cdot p_Y(y)$$

- Continuous random variables X and Y are said to be independent if

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

- If these conditions don't hold then X and Y are said to be dependent.

Expected value

- The expected value of a function $h(x, y)$, denoted $E[h(X, Y)]$, is defined as

$$E[h(X,Y)] = \begin{cases} \sum_x \sum_y h(x,y) \cdot p(x,y) & \text{discrete} \\ \int_{-\infty}^{\infty} \int_{\infty}^{\infty} h(x,y) \cdot f(x,y)\, dx\, dy & \text{continuous} \end{cases}$$

- The covariance between X and Y is defined as

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= \begin{cases} \sum_x \sum_y (x-\mu_X)(y-\mu_Y) p(x,y) & \text{discrete} \\ \int_{-\infty}^{\infty} \int_{\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f(x,y)\, dx\, dy & \text{continuous} \end{cases} \end{aligned}$$

- Sometimes it is more convenient to evaluate $\mathrm{Cov}(X,Y) = E[XY] - \mu_X \mu_Y$

Correlation

- The correlation coefficient of X and Y, denoted $\text{Corr}(X, Y)$ or $\rho_{X,Y}$ or simply ρ, is defined as

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

- For any two rv's X and Y, $-1 \leq \rho_{X,Y} \leq 1$
- If a and c are either both positive or both negative then

$$\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$$

- If X and Y are independent, then $\rho = 0$. However, $\rho = 0$ does not imply that X and Y are independent.
- $\rho = -1$ or $\rho = 1$ if and only if $Y = aX + b$ for some numbers a and b.

Overview

- **Probability**
  - Random variables and probability
  - Discrete distributions
  - Continuous distributions
  - Joint probability of multiple random variables

- **Statistics**
  - Sampling
  - Statistical inference
  - Estimation
  - Hypothesis testing

Random samples

- Evaluating the distribution of a statistic calculated from a sample with an arbitrary joint distribution can be very difficult.
- Frequently we make the simplifying assumption that our data constitute a random sample $X_1, X_2, \ldots, X_n$ from a distribution. This means that
  1. The X_i's are independent.
  2. All the X_is have the same probability distribution

Linear Combinations and their means

- Given a collection of n random variables $X_1, X_2, \ldots, X_n$ and n numerical constants $a_1, a_2, \ldots, a_n$, the random variable

$$Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

  is called a linear combination of the X_is.
- Whether or not the X_is are independent,

$$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n]$$

Variances of linear combinations

- If $X_1, X_2, \ldots, X_n$ are independent with variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ then

$$\begin{aligned} V(a_1 X_1 + a_x X_2 + \cdots + a_n X_n) & \\ &= a_1^2 V(X_1) + a_2^2 V(X_2) + \cdots + a_n^2 V(X_n) \\ &= a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \cdots + a_n^2 \sigma_n^2 \end{aligned}$$

- In general

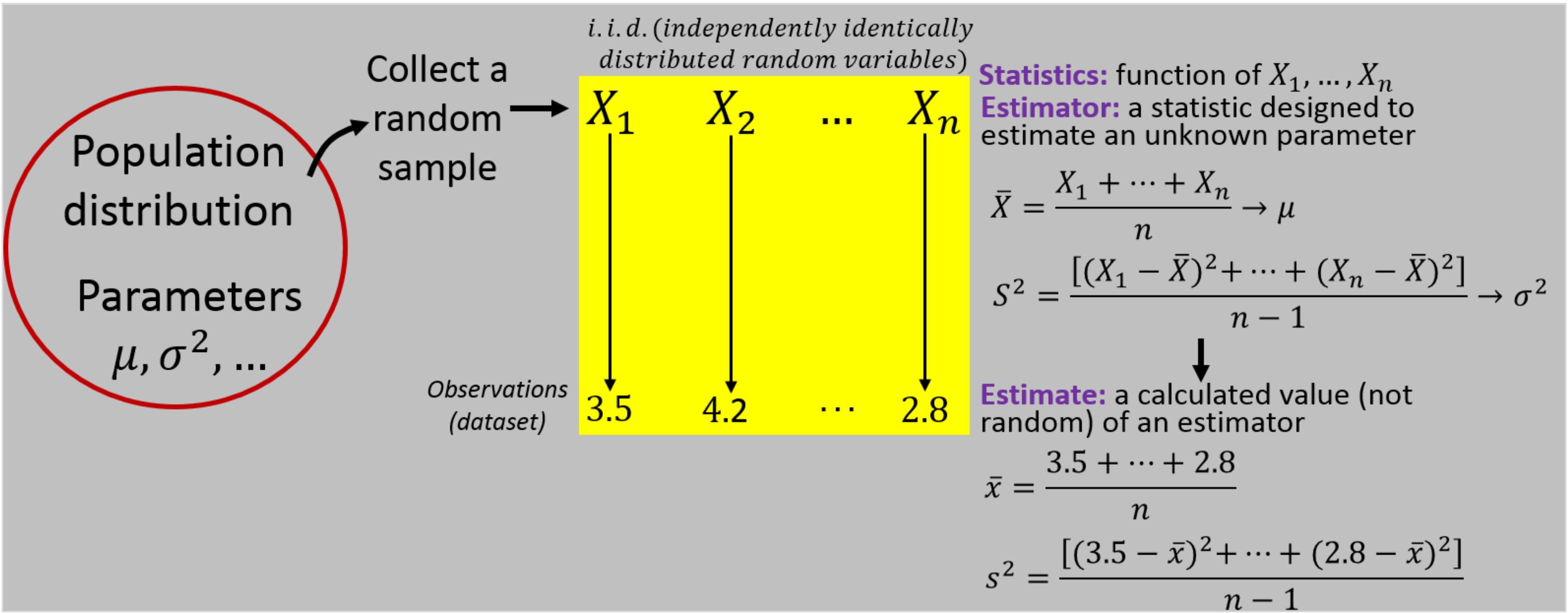
$$V(a_1 X_1 + a_x X_2 + \cdots + a_n X_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \mathrm{Cov}(X_i, X_j)$$

The Case of Normal Random Variables

- When the X_is are independent and normally distributed, any linear combination will also be normally distributed.

Statistical Inference

Statistical inference: Find truth on the population based on the data obtained from a sample of the population



- **Estimation:** Find estimates of the unknown parameters
  - Point estimation: $\hat{\mu} = 2.5$
  - Confidence interval (CI) estimation: the 95% CI of μ = $(2.0, 3.0)$
- **Hypothesis testing:** Decisions based on specific hypotheses (e.g., $\mu \leq 2$ $vs.$ $\mu > 2$)

Point Estimation

- A **point estimator** is designed to estimate an unknown parameter with a single value

  θ = unknown parameter

  $\hat{\theta}$ = point estimator (a function of the data)

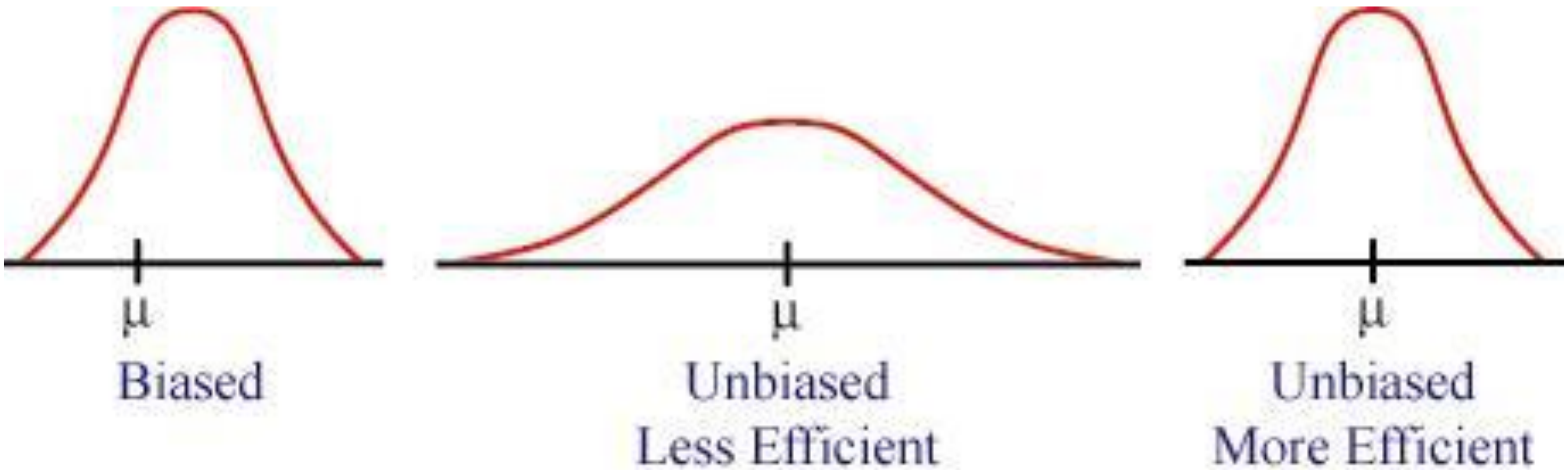
- **Example:** $\hat{\mu} = \bar{X}$ estimates μ

- **How do we identify a good point estimator?**
  - An estimator $\hat{\theta}$ is **unbiased** iff $E\left(\hat{\theta}\right) = \theta$
  - If an estimator $\hat{\theta}$ has the smallest variance, then it is the **most efficient** estimator of θ

Example Sampling Distribution of $\hat{\boldsymbol{\theta}}$

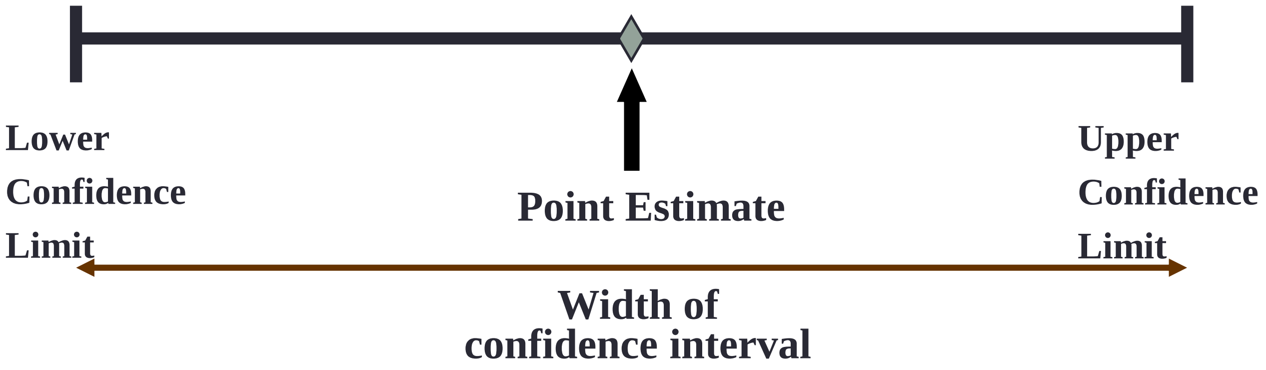
$\theta = \mu$

Red curve: Distribution of $\hat{\mu}$



Confidence Interval Estimation

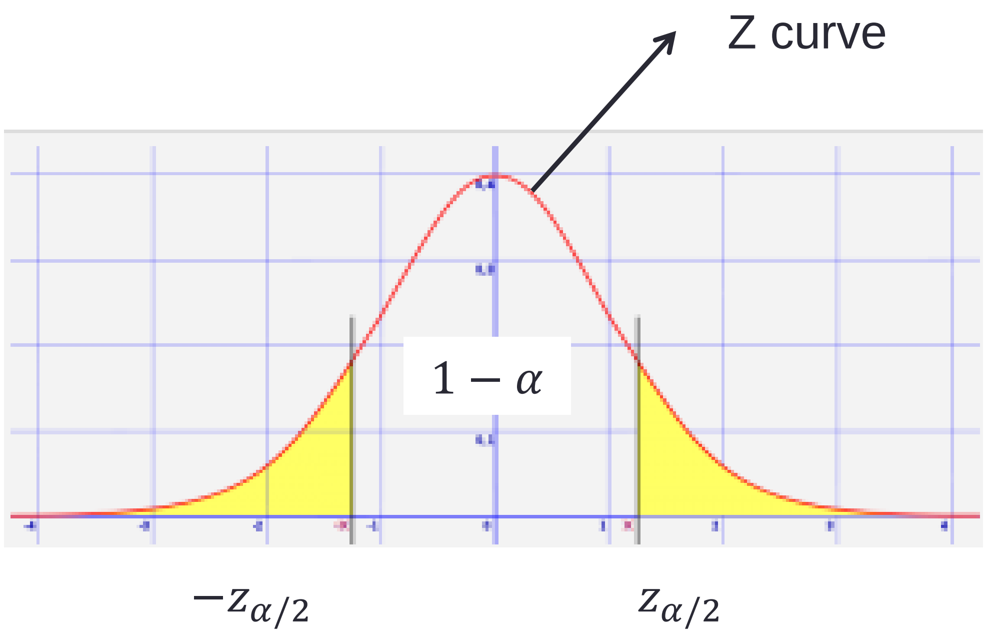
- **Interval estimate:** an entire interval of plausible values
  - More information about a population than does a point estimate
  - A confidence level for the estimate
- **Confidence level:** a measure of degree of reliability of the interval (95%, 99%, 90%)
- **Significance level (α):** 1 − confidence level
- **Width of CI:** given the confidence level, if the interval is narrow, our knowledge of the parameters is reasonably precise; a very wide CI indicates large amount of uncertainty.



CI of Normal Distribution

- A $100(1-\alpha)\%$ confidence interval for the mean μ of a normal population **when the value of σ is known** is given by

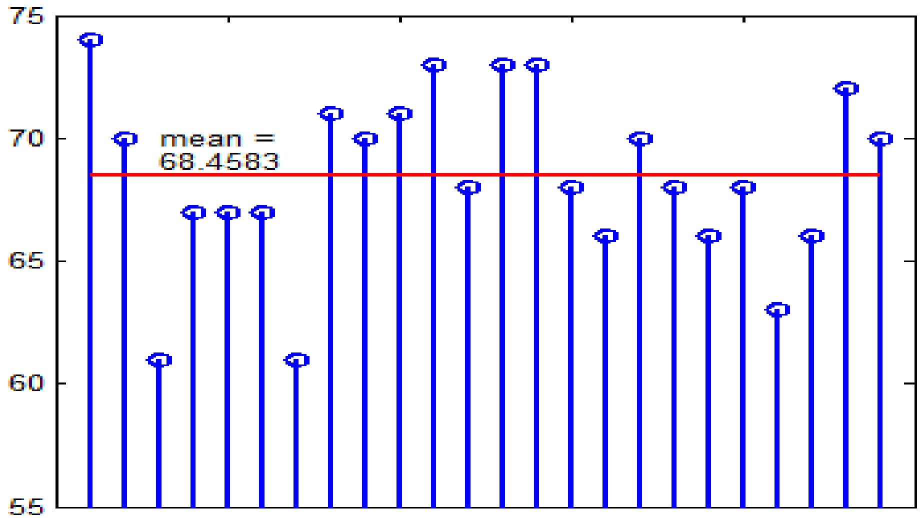
$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$



$$\text{P}(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$$

Example

- **University student height:** given $n = 24, \bar{x} = 68.46, \sigma = 2$
- 95% confidence interval: $\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}},\ \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right)$
- $\left(68.46 - 1.96\frac{2}{\sqrt{24}}, 68.46 + 1.96\frac{2}{\sqrt{24}}\right) = (67.66, 69.26)$



CI When Variance Unknown

- **Assumption:** population is normal, and random samples are from a normal distribution with both μ and σ unknown.

- Let $\bar{x}$ and s be the sample mean and sample standard deviation from a normal population with mean μ. Then the $100(1-\alpha)\%$ confidence interval for μ is

$$\left(\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right)$$

- **Critical value:** Let $t_{\alpha,\nu}$ denote the number on the measurement axis for which the area under the *t* curve with ν DoF to the right of $t_{\alpha,\nu}$ is α; $t_{\alpha,\nu}$ is called a *t* critical value.

Hypothesis Testing

- **Hypothesis test:** a method of making decisions using data, whether from a controlled experiment or an observation study (not controlled), that produces a conclusion about the population
  - Example: Is there a difference between the accuracy of two gauges based on sample data?
  - The problem conjecture is put in the form of **statistical hypothesis**
  - Rejection/non-rejection of the hypothesis is made using statistical inference procedure
- **Statistical hypothesis:** an assertion or conjecture concerning one or more populations.
- **Performance**
  - **Type I error (α):** rejection of the null hypothesis when it is true
  - **Type II error (β):** non-rejection of the null hypothesis when it is false

Procedure

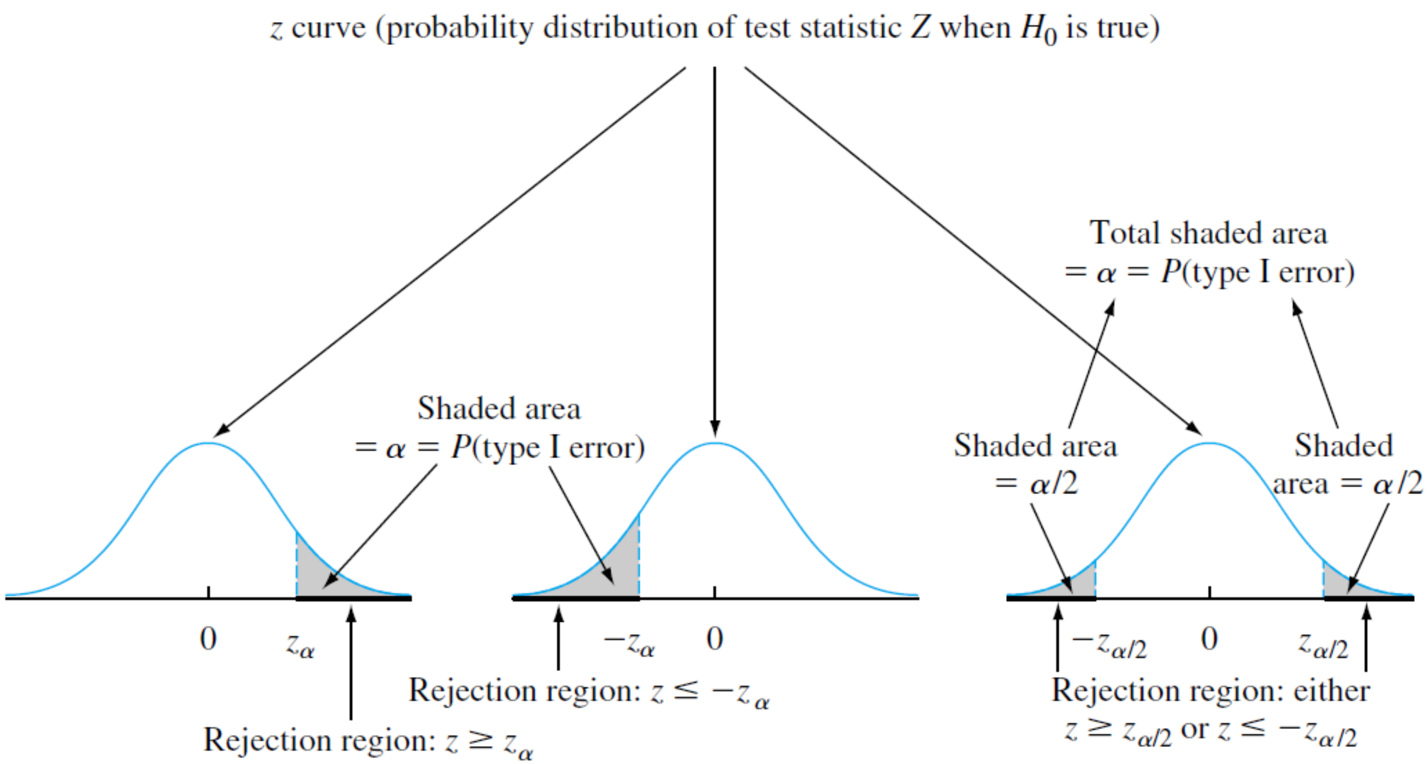
1. State the null hypothesis (H_0)

   "nothing" hypothesis, nothing has changed, of no difference, nothing special taking place, no systematic effect
2. State alternative hypothesis (H_a)

   Researcher's conjecture, paranoia, change, effect of treatment
3. Choose the test statistic (e.g., z vs. t for mean)
4. Determine the critical value and rejection region
5. Calculate the test statistic value
6. Reject H_0 if the test statistic is within the critical region or *p*-value $<\alpha$; otherwise, do not reject
7. Draw the conclusions/implications

Critical Value and Rejection Region

H_a: $\mu > \mu_0$	$z \geq z_\alpha$	(upper-tailed test)
H_a: $\mu < \mu_0$	$z \leq -z_\alpha$	(lower-tailed test)
H_a: $\mu \neq \mu_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$	(two-tailed test)



Type I and Type II Errors

- **Type I Error**
  - If we reject H_0 when in fact H_0 is true. This would be akin to convicting an innocent person for a crime(s) he did not commit.

- **Type II Error**
  - If we fail to reject H_0 when in fact H_a is true. This is analogous to a guilty person escaping conviction.

- **Type I errors are usually considered worse**, so we design our statistical procedures to control the probability of making such a mistake. We define the

$$significance\ level\ of\ the\ test = P(Type\ I\ error) = \alpha$$

Significance Level

- We want α to be small which conventionally means, say, $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.005$
- **Rejection region** for a test is the set of sample values which would result in the rejection of H_0
  - For previous example, the rejection region would be all possible samples that result in a 95% confidence interval that does not cover $\mu = 70$.
- The above example with H_a: $\mu \neq 70$ is called a **two-sided test**. Sometimes we are interested in a one-sided test, which would look like H_a: $\mu < 70$ or H_a: $\mu > 70$.

P Value

- **P value:** the lowest level (of significance) at which the observed value of the test statistic is significant

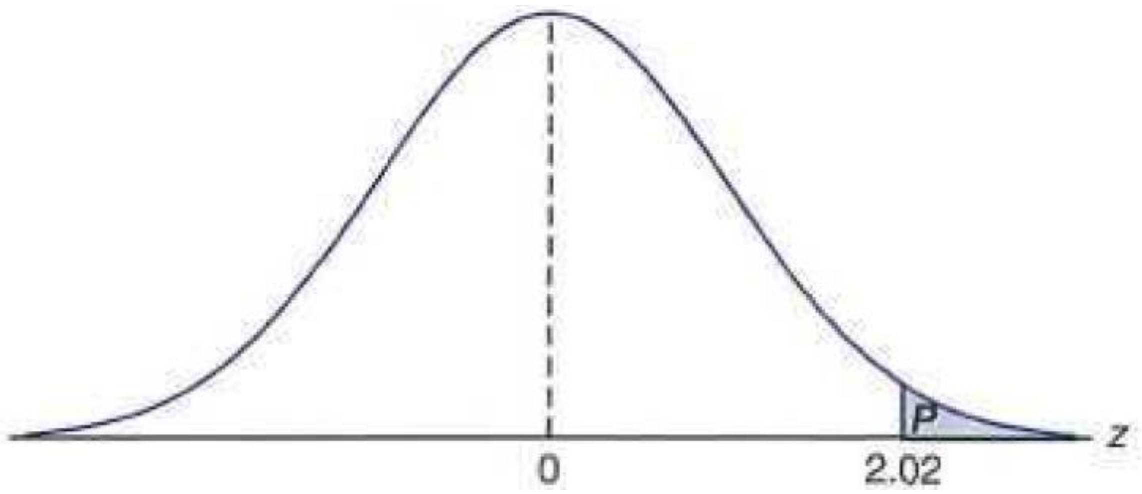
- The plausibility of the null hypothesis H_0

Example

- A random sample of machines in a plant showed an average useful life of 71.8 months. Assuming a population standard deviation of 8.9 months, does this seem to indicate the mean useful life is greater than 70 months?
- Solution
  - H_0: $\mu = 70$
  - H_1: $\mu > 70$
  - $\alpha = 0.05$, test statistic $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$
  - Rejection region: $z > 1.645$
  - Test statistic: $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{71.8-70}{8.9/\sqrt{100}} = 2.02 > 1.645$
  - Reject H_0 at $\alpha = 0.05$
  - Conclusion: there is significant evidence that the mean useful life is greater than 70 months.

P-value Solution

- $p = P(z > 2.02) = 0.0217 < 0.05$
- As a result, the evidence in favor of H_1 is stronger than that suggested by a 0.05 level of significance. That means there is significant evidence that the mean useful life is greater than 70 months.



Popular Tests

One sample

- For mean: z-test (large sample size or normal population with σ known), t-test (small sample of normal population with σ unknown)
- For variance: χ^2-test (normal population)

Two sample

- For mean: z-test, t-test
- For variance: F-test

Multivariate (one sample):

- For mean: T^2-test