

Statistical Hypothesis:

- A statistical hypothesis is a statement about the parameters of one or more populations.
- Since we use probability distributions, a statistical hypothesis may be a statement about the probability distribution of a random variable. Indeed, the hypothesis will usually involve one or more parameters of this distribution.

Null Hypothesis:

- In many instances we formulate a statistical hypothesis for the purpose of rejecting or nullifying it.

Example 1: Suppose you want to decide whether a coin is fair. If you toss it 100 times and get 85 heads, would you think the coin is likely to be unfair? What about 60 heads? Or 52 heads?

- Most people would guess that 85 heads is a strong evidence that the coin is unfair. However, 52 heads no evidence at all. Sixty heads is less clear.
- Null hypothesis significance testing (NHST) is a method (frequentist) to thinking quantitatively about these questions.

Ingredients for NHST:

H_0 : the null hypothesis.

H_A : The alternative hypothesis:

If we reject the null hypothesis we accept this alternative as the best explanation for the data.

X : the test statistic.

We compute this from the data.

Null distribution: The probability distribution of X assuming H_0 .

Rejection Region: If X is in the region we reject H_0 in favor of H_A .

Non-Rejection Region: This is the complement of the rejection region. If X is in the region, we do not reject H_0 .

Example 2:

- To test whether a coin is fair we flip it 10 times. If we get an unexpectedly large or small number of heads we will suspect the coin is unfair.
- Let θ be the probability that the coin lands heads when flipped.

1. Null hypothesis:

H_0 = "the coin is fair", namely, $\theta = 0.5$.

2. Alternative hypothesis:

H_A = "the coin is not fair, namely, $\theta \neq 0.5$ ".

3. Test statistic:

X = number of heads in 10 flips.

4. Null distribution:

- This is the probability function based on the null hypothesis.

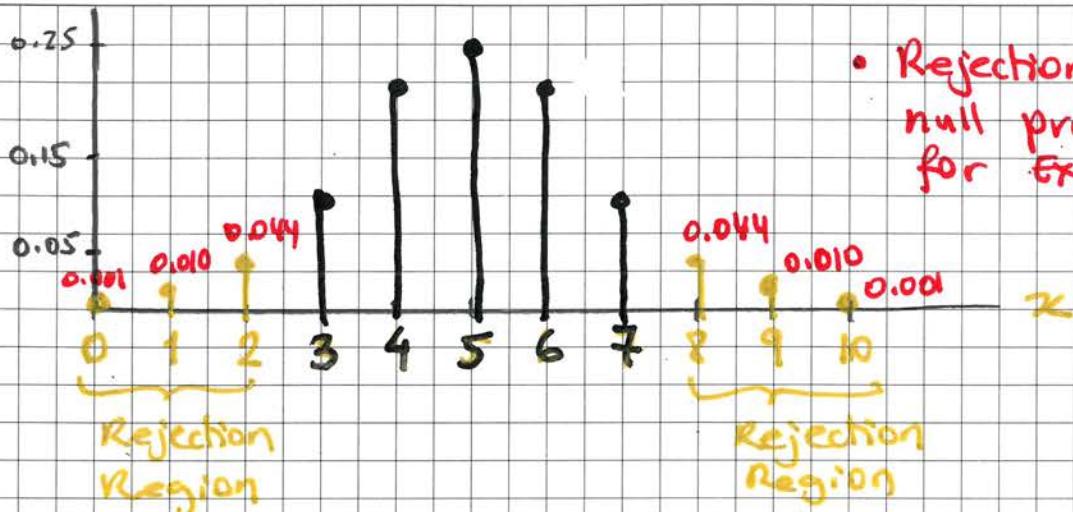
$$P(x | \theta=0.5) \sim \text{Binom}(10, 0.5).$$

- Here is the probability table for the null distribution.

x	0	1	2	3	4	5	6
$P(x H_0)$	0.001	0.10	0.044	0.117	0.205	0.246	0.205
x	7	8	9	10			
$P(x H_0)$	0.117	0.044	0.010	0.001			

5. Rejection region:

- Under the null hypothesis we expect to get about 5 heads in 10 tosses.
- We will reject H_0 if the number of heads is much fewer or greater than 5.
- Let Set the rejection region as $\{0, 1, 2, 8, 9, 10\}$. If the number of heads in 10 tosses is in this region we will reject the hypothesis that the coin is fair.



- Rejection region and null probabilities for Example 2.

Remarks For Example 2:

- We won't claim the coin is unfair unless we have compelling (zorlayıcı) evidence.
- The rejection region consists of data that is extreme under the null hypothesis. They are in the tail of the null distribution away from the high probability center.
 - How far away depends on the significance level α of the test. We will discuss soon!
- If we get 3 heads in 10 tosses, then the test statistic is in the non-rejection region.
"Data does not support rejecting the null hypothesis".
- Even if we got 5 heads, we would not claim that the data proves the null hypothesis is true.

Question: If we have a fair coin what is the probability that we will decide incorrectly it is unfair?

Answer: The null hypothesis is that the coin is fair. The question asks for the probability the data from a fair coin will be in the rejection region. This is the sum of the probabilities {0, 1, 2, 8, 9, 10} heads in 10 tosses. Answer is the sum of these probabilities, namely,

$$P_0 + P_1 + P_2 + P_8 + P_9 + P_{10} = 0.11.$$

We can write,

$$\Pr \{ \text{rejecting } H_0 \mid H_0 \text{ is true} \} = 0.11.$$

Using Example 2, we will continue to define more terms used in NHST and see how to quantify properties of significance test.

- Definition:

SIMPLE HYPOTHESIS: A simple hypothesis is one for which we can specify its distribution completely.

- For example, the null hypothesis in Example 2 is that $\theta = 0.5$, so the null distribution is binomial(10, 0.5). Since the null distribution is fully specified, H_0 is simple.

- Definition:

COMPOSITE HYPOTHESIS: If its distribution cannot be specified fully, we say that the hypothesis is composite.

- A parameter of interest lies in a range of values for a typical composite hypothesis.
- Indeed, in Example 2 the alternative hypothesis is that $\theta \neq 0.5$. This is really many hypotheses in one. θ could be 0.51, 0.7, 0.99, etc.
- Since the alternative distribution binomial(10, θ) is not fully specified, H_A is composite.

Example 3:

Date.

No. 57

- Suppose we have data x_1, \dots, x_n .
Suppose our hypotheses are
 H_0 : the data is drawn from $N(0, 1)$,
 H_A : the data is drawn from $N(1, 1)$.
- These are both simple hypotheses: each hypothesis completely specifies the distribution.

Example 4: (Composite hypothesis.)

- Now suppose that our hypothesis are

H_0 : the data is drawn from a Poisson distribution of unknown parameter.

H_A : the data is not drawn from a Poisson distribution.

- These are both composite hypotheses, indeed, they don't fully specify the distribution.

Example 5: In an Extrasensory perception or ESP experiment (Sixth sense), a subject is asked to identify the suits (diamonds, spades, clubs, or hearts) of 100 cards drawn with replacement from a deck cards.

- Let T be the number of successes.
- The (simple) null hypothesis that the subject does not have ESP is given by
 $H_0 : T \sim \text{binomial}(100, 0.25)$

- The Composite or alternative hypothesis that the subject has ESP is given by

$$H_A : T \sim \text{binomial}(100, p) \text{ with } p > 0.25$$

- Another composite or alternative hypothesis that something besides pure chance is going on

$$H_A : T \sim \text{binomial}(100, p) \text{ with } p \neq 0.25$$

In this case, the subject has ESP or anti-ESP. Indeed, values of $p < 0.25$ represent hypotheses that the subject has a kind of ANTI-ESP.

TYPES OF ERRORS

Type I Error: (false rejection of H_0)

- Rejecting the null hypothesis H_0 when it is true is defined as type I error.

Type II Error: (false non-rejection ("acceptance") of H_0)

- Failing to reject the null hypothesis when it is false is defined as type II error.

- These situations are presented in the following table:

Decisions in Hypothesis Testing

Date.

No. 59

<u>Our Decision</u>	H_0 is True	H_A is true [H_0 is false]
Reject H_0	Type I error	No error [Correct decision]
Don't reject H_0 (fail to reject)	No error (Correct decision)	Type II error

Probability of Making Type-I Error:

$$\alpha = \Pr\{\text{Type-I error}\} = \Pr\{\text{reject } H_0 \text{ when } H_0 \text{ is true}\}$$

- Sometimes the type-I error probability is called the **Significance level**, or the α -error.

Probability of Making Type-II Error:

$$\beta = \Pr\{\text{Type-II error}\} = \Pr\{\text{fail to reject } H_0 \text{ when } H_0 \text{ is false}\}$$

- To calculate β , we must have a specific alternative hypothesis.

Power:

- The **power** of a statistical test is the probability of rejecting the null hypothesis H_0 when the alternative hypothesis is true.

• Significance level and Power

are used to quantify the quality of the significance test.

- Ideally a significance test would not make errors. Another words, it would not reject H_0 when H_0 was true and would reject H_0 in favor of H_A when H_A is true.
- In the above table about decisions in hypothesis testing, There are 4 important probabilities:

$$\Pr\{\text{reject } H_0 \mid H_0 \text{ true}\}$$

$$\Pr\{\text{reject } H_0 \mid H_A \text{ true}\}$$

$$\Pr\{\text{do not reject } H_0 \mid H_0 \text{ true}\}$$

$$\Pr\{\text{do not reject } H_0 \mid H_A \text{ true}\}$$

- The two probabilities we focus on are:

$$\text{Significance level} = \Pr\{\text{reject } H_0 \mid H_0 \text{ true}\}$$

= probability we incorrectly reject H_0
= $\Pr\{\text{Type I error}\}$

$$\text{Power} = \text{probability we correctly reject } H_0$$

$$= \Pr\{\text{reject } H_0 \mid H_A \text{ true}\}$$

$$= 1 - \Pr\{\text{Type II error}\}$$

Remarks:

Date.

No. 61

- Ideally, a hypothesis test should have a small significance level (near zero) and a large power (near one).
- The following are two analogies to help you remember the meaning of significance level and power:

(1) Think of H_0 as the hypothesis:

- "nothing noteworthy is going on", namely,
- "The coin is fair",
- "The treatment is no better than placebo",

And think of H_A as the opposite:

- "Something interesting happening"

- Then power is the probability of detecting something interesting when it is present and significance level is the mistakenly claiming something interesting has occurred.

(2) In the U.S. Criminal defendants are presumed innocent until proven guilty beyond a reasonable doubt. We can state this in terms of NHST as

H_0 : the defendant is innocent (the default)

H_A : the defendant is guilty.

- The power is the probability of correctly finding a guilty party guilty. Significance level is the probability of finding an innocent person guilty.

"Beyond a reasonable doubt" means we should demand the significance level be very small."

COMPOSITE HYPOTHESIS

- In Example 2, alternative hypothesis is composite because

$$H_0 : \theta = 0.5 \quad (\text{Simple hypothesis, fully described})$$

$$H_0 \sim \text{binom}(10, 0.5)$$

$$H_A : \theta \neq 0.5 \quad (\text{Not fully described, it could be } 0.51, 0.7, 0.99)$$

$$H_A \sim \text{binom}(10, \theta)$$

Therefore, the power is different for different values of θ . We look at likelihoods. The probability of the data given a hypothesis.

x	0	1	2	3	4	5	6	7	8	9	10
$H_0 : p(x \theta = 0.5)$.001	.010	.044	.117	.205	.246	.205	.117	.044	.010	.001
$H_A : p(x \theta = 0.6)$.000	.002	.011	.042	.111	.201	.251	.215	.121	.040	.006
$H_A : p(x \theta = 0.7)$.000	.0001	.001	.009	.037	.103	.200	.267	.233	.121	.028

- We can compute the significance level and the power of this test;
- Significance level = Probability we reject H_0 when it is true.
 = Probability the test statistic is in the rejection region, namely $\{0, 1, 2, 8, 9, 10\}$, when H_0 is true.
 = Sum of the rejection region in the $\theta=0.5$ row.
 = 0.11

Power

when $\theta=0.6$ = Probability we reject H_0 when $\theta=0.6$.

= Probability the test statistic is in the rejection region when $\theta=0.6$ row.

= $P_0 + P_1 + P_2 + P_8 + P_9 + P_{10}$, probability of the test statistic is in the rejection region of the $\theta=0.6$ row of the table

$$= 0.180,$$

POWER

when $\theta=0.7$ = Probability we reject H_0 when $\theta=0.7$.

= Probability the test statistic is in the rejection region when the $\theta=0.7$ row of the table.

$$= 0.384.$$

Remarks:

- We see that the power is greater for $\theta=0.7$ than for $\theta=0.6$. This is not surprising since we get higher power when alternate hypothesis is farther from the null hypothesis.
- It should be quite hard to distinguish a fair coin from one with $\theta=0.51$.
- Power is a very descriptive and concise measure of the sensitivity of a statistical test, where by sensitivity we mean the ability of the test to detect differences.

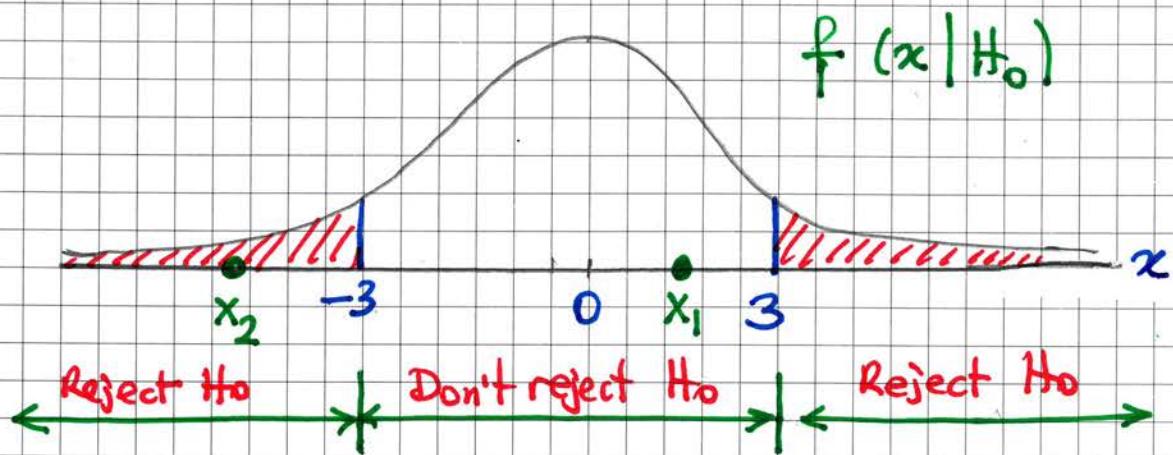
CONCEPTUAL SKETCHES

Date.

No. 64

1) Null Distribution: Rejection and Non-rejection regions

- The following diagram below shows a null distribution with rejection and non-rejection regions. x_1 and x_2 two possible test statistics.



- The test statistic x_1 is in the non-reject region. Therefore, if our data produced test statistic x_1 , then we would not reject the null hypothesis H_0 .
- However, the test statistic x_2 is in the rejection region, so if our data produced x_2 , then we would reject the null hypothesis in favor of the alternative hypothesis.

OBSERVATIONS:

- The rejection region consists of values far from the center of the null distribution.
- The rejection region is two-sided. There are examples of one-sided rejection regions as well.

(3) The alternative hypothesis is not mentioned. No. 65

We reject or don't reject H_0 based only on the likelihood $f(x|H_0)$, namely, the probability of the test statistic conditioned on H_0 .

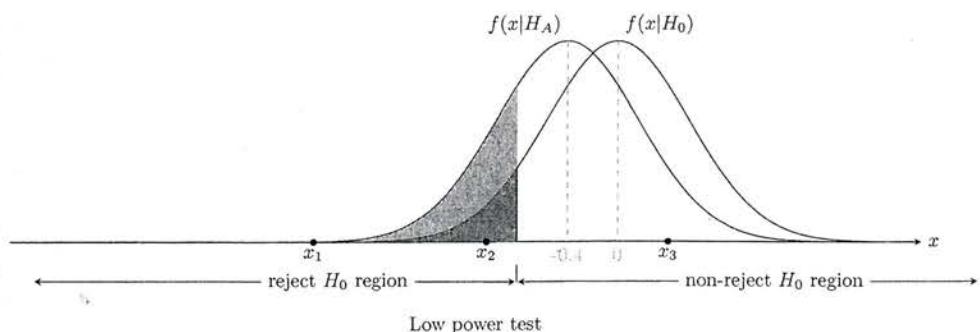
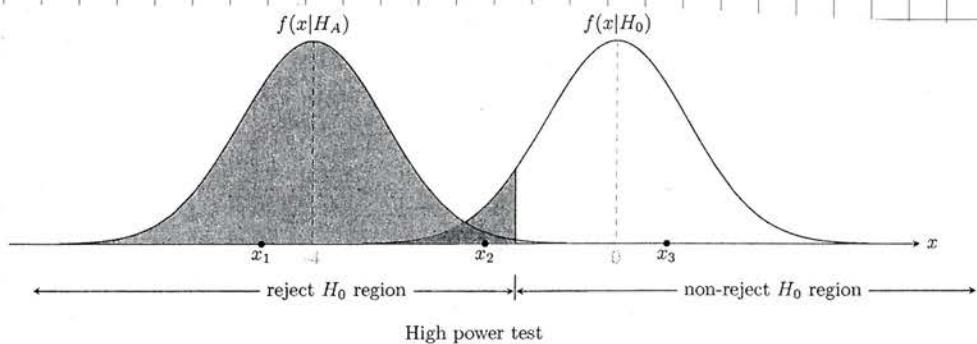
- The alternative hypothesis H_A should be considered when choosing a rejection region, but formally it does not play a role in rejecting or non-rejection H_0 .

(4) Sometimes we call the non-rejection region the "acceptance region". This is technically incorrect because we never truly accept the null hypothesis.

! WARNING! !

- You can never prove the null hypothesis, we either reject or say the data does not support rejecting H_0 .

2. HIGH and LOW POWER TEST



- The above two figures show high and low power tests:
 - The shaded area under $f(x|H_0)$ represents the significance level:
 - The probability of falsely rejecting the null hypothesis when it is true. or
 - The probability the test statistic falls in the rejection region even though H_0 is true.
- Similarly, the shaded area under $f(x|H_A)$ represents the power:
 - Both tests have the same significance level.
 - If $f(x|H_A)$ has considerable overlap with $f(x|H_0)$, the power is much lower.
- In both tests both distributions are standard normal.
 - The null distribution, rejection region and significance level are all the same.
 - In the top figure we see that the means of two distributions are 4 standard deviations apart. Since the areas under the densities have very little overlap, the test has high power.
 - This means that if the data x drawn from H_A , it will almost certainly be in the rejection region. x_3 is very surprising outcome for the H_A distribution.

- In the bottom figure we see the means of the two distributions are just 0.4 standard deviations apart.
- Since the areas under the densities have a lot of overlap, the test has a low power.
- If the data x drawn from H_A it is likely to be in the non-rejection region. Therefore, x_3 would not be very surprising outcome for the H_A distribution.
- Normally, we can increase the power of a test by increasing the amount of data and hence decreasing the variance of the null and alternative distributions.

Example 6:

- Suppose a drug for a disease is being compared to be placebo. We choose our null and alternative hypotheses as :

H_0 : The drug does not work better than the placebo.

H_A : The drug works better than the placebo.

- The power of hypothesis test is the probability that the test will conclude that the drug is better, if it is indeed truly better.
- The significance level is the probability that the test will conclude that the drug works better, when in fact it does not, $\Pr \{ \text{reject } H_0 \mid H_0 \text{ true} \}$, work better than the placebo.

DESIGNING a HYPOTHESIS TEST

1) Pick the null hypothesis H_0 :

- The choice of H_0 and H_A is not mathematics. It is art and custom. We choose H_0 to be simple, namely,
 - No effect of drug,
 - No extrasensory perception,
 - No bias in the coin.

2) Decide if H_A is one-sided or two-sided:

- For example, in Example 2 we wanted to know if the coin is fair.
 - An unfair coin could be biased for or against heads, so

$$H_A : \theta \neq 0.5$$

is a two-sided hypothesis If we only care whether or not the coin is biased

for heads we could use the one-sided hypothesis:

$$H_A : \theta > 0.5.$$

3. Pick a Test Statistic:

- This test statistic can be the sample mean, sample total, or sample variance. The distributions that go with statistics are always conditioned on the null hypothesis.
- Some standard statistics that we will encounter are Z , t , and χ^2 .
- We will compute likelihoods such as $f(z | H_0)$.

4. Pick a Significance Level and Determine the Rejection Region

- We use α to denote the significance level, we pick α in advance. Typical values are 0.1, 0.05, 0.01.
- Remember that the significance level is the probability of type-I error, namely incorrectly rejecting hypothesis when it is true.

- Once the significance level is chosen we can determine the rejection region in tail(s) of the null distribution.
- In Example 2, H_A is two-sided so the rejection is split between the two tails of the null distribution. This distribution is given in the following table.

x	0	1	2	3	4	5	6	7	8	9	10
$p(x H_0)$.001	.010	.044	.117	.205	.246	.205	.117	.044	.010	.001

- If we set $\alpha=0.05$, then the rejection region must contain at most 0.05 probability
 - For a two-sided rejection we get
Rejection Region = {0, 1, 9, 10}.
 - If we set $\alpha=0.01$, the rejection region is
 $\{0, 10\}$.
 - Suppose we change H_A to
 H_A ; "the coin is biased" in favor of heads.
We now have one-sided hypothesis
 $H_A: \theta > 0.5$.

- Now our rejection will be in the right-hand tail. Because we do not want to reject H_0 in favor of H_A . If we get a smaller number of heads.
- Now if $\alpha = 0.05$, the rejection region is the one-sided range $\{9, 10\}$.
- If we set $\alpha = 0.01$ the rejection region is $\{10\}$.

5. Determine the Power(s):

- As shown in Example 2, once the rejection is set, we can determine the power test as various values of the alternate hypothesis.

Consequences of Significance:

- If $\alpha = 0.1$, then we would expect a 10% type I error rate. In other words, we expect to reject the null hypothesis in 10% of those experiments when the null hypothesis is true.
- For example, if you were running an experiment to determine if your chocolate is more than 72% cocoa, then a 10% type I error rate is fine.

- However, if you are identifying fingerprints for a murder trial, then a 10% type I error rate is definitely not acceptable. This means that we are mistakenly claiming that fingerprints found at the crime scene belonged to someone who is truly innocent.

SIGNIFICANCE FOR A COMPOSITE NULL HYPOTHESIS

- If H_0 is composite then $\Pr\{\text{type I error}\}$ depends on which member of H_0 is true. In this case the significance level is defined as the maximum of these probabilities.

CRITICAL VALUES:

- The sample median is measure of central tendency that divides the data into equal parts, half below and half above.
- We can also divide data into more than two parts. When the ordered set of data is divided into four equal parts, the division points are called quartiles.

Definition: P-probability

- A p-probability interval for θ is an interval $[a, b]$ with $\Pr\{a \leq \theta \leq b\} = p$.

(1) In the continuous case with pdf $f(\theta)$, this means

$$\int_a^b f(\theta) d\theta = p.$$

(2) In the discrete case with pmf $p(\theta)$, this means

$$\sum_{a \leq \theta_i \leq b} p(\theta_i) = p.$$

(3) We may say 90% probability interval to mean 0.9-probability interval.

- Probability intervals are also called credible intervals to contrast them with confidence intervals.

Example:

- Between the 0.05 and 0.55 quantiles is a 0.5 probability interval. There are many 50% probability intervals. In particular, notice that p-probability interval for θ is not unique.

Q - Notation:

- We can explain probability intervals in terms of quantiles:

- s -quantile for θ is the value q_s with

$$\Pr\{\theta \leq q_s\} = s.$$

- Therefore, for $s \leq t$, the amount of probability between the s -quantile and t -quantile is just $t-s$. In these terms, a p -probability interval $[q_s, q_t]$ with $t-s=p$.

Remarks:

- For example, we may have 0.5-probability intervals

$$[q_{0.25}, q_{0.75}] \text{ and } [q_{0.05}, q_{0.95}].$$

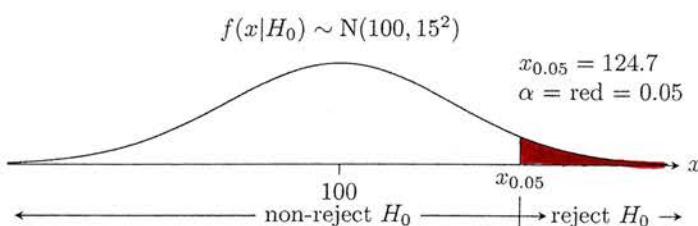
- The interval $[q_{0.25}, q_{0.75}]$ is symmetric because the amount of probability remaining on either side of the interval is the same, namely, 0.25.

- Different p -probability intervals for θ may have different widths.

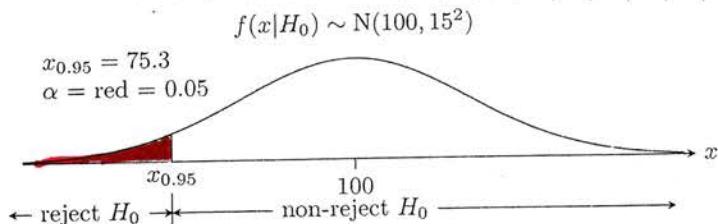
Example:

- (1) Let's find the 0.05 critical value for the standard normal distribution.

- We label this critical value $Z_{0.05}$.
- The critical value $Z_{0.05}$ is just the 0.95 quantile. This means, it has 5% probability to its right and therefore 95% probability to its left.
- We compute $Z_{0.05} = 1.64$. $Z \sim N(0,1)$.
- The value of the significance test that marks the start of the rejection region is the critical value. In the above example $Z_{0.05} = 1.64$ is a critical value.

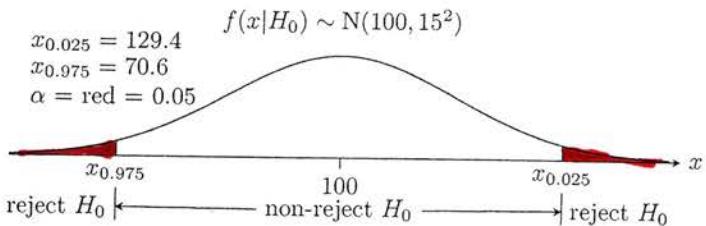


- (2) Repeat the previous example for a left sided rejection region with a significance level $\alpha=0.05$.



Since there is a 95% probability to its right, the start of the rejection is at the 0.05 quantile. So we label it $x_{0.95}$, we compute it as $x_{0.95} = 75.3$.

- (c) Repeat the previous example for a two-sided rejection region. Put half the significance in each tail.



- We have a total significance of 0.05. We put 0.025 in each tail.
- The left tail starts at $x_{0.975} = 70.6$ and the right tail starts at $x_{0.025} = 129.4$.

LARGE-SAMPLE CONFIDENCE INTERVAL FOR μ

- Suppose that the population distribution is normal with unknown mean and known standard deviation σ .
- Large-Sample Confidence interval for μ does not require these assumptions.
- Let X_1, X_2, \dots, X_n be the random sample from a population with an unknown mean μ and unknown variance σ^2 .
- If the sample size n is large, the central limit theorem implies that \bar{X} has approximately a normal distribution with mean μ and

variance $\frac{\sigma^2}{n}$. Therefore,

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

has approximately a standard normal distribution.

- However, the standard deviation σ is unknown. When n is large, replacing σ by the sample standard deviation S' has a little effect on the distribution of Z .
- This leads to the following useful result.

LARGE - SAMPLE CONFIDENCE INTERVAL on the MEAN

($n > 30$)

- When n is large, the quantity

$$\frac{\bar{X} - \mu}{\frac{S'}{\sqrt{n}}}$$

has an approximate standard normal distribution. Consequently,

$$\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}} \quad (*)$$

is a large sample confidence interval (CI) for μ , with confidence level approximately $100(1-\alpha)\%$.

- Equation (*) holds regardless of the shape of the population distribution. Generally n should be at least 40 to use this result reliably.
- The central limit theorem generally holds for $n \geq 30$, but larger sample size is recommended here because replacing σ by s in Z results in additional variability.

Example: [Montgomery]

- A sample fish selected from 53 florida lakes and merkury concentration in the muscle tissue was measured.

1.230	0.490	0.490	1.080	0.590	0.280	0.180	0.100	0.940
1.330	0.190	1.160	0.980	0.340	0.340	0.190	0.210	0.400
0.040	0.830	0.050	0.630	0.340	0.750	0.040	0.860	0.430
0.044	0.810	0.150	0.560	0.840	0.870	0.490	0.520	0.250
1.200	0.710	0.190	0.410	0.500	0.560	1.100	0.650	0.270
0.270	0.500	0.770	0.730	0.340	0.170	0.160	0.270	

- We want to find an approximate 95% CI on μ . Because $n > 40$, the assumption of normality is not necessary to use equation (*). The required quantities are

$$n = 53$$

$$\bar{x} = 0.5250$$

$$s = 0.3486$$

$$z_{0.025} = 1.96$$

The approximate 95 % CI on μ is

$$\bar{x} - z_{0.025} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{0.025} \frac{s}{\sqrt{n}}$$

$$0.5250 - 1.96 \frac{0.3486}{\sqrt{53}} \leq \mu \leq 0.5250 + 1.96 \frac{0.3486}{\sqrt{53}}$$

$$0.4311 \leq \mu \leq 0.6189 .$$

This interval is fairly wide because there is a lot of variability in mercury measurements.

CONFIDENCE INTERVAL (IC) ON THE MEAN OF NORMAL DISTRIBUTION, VARIANCE UNKNOWN

- When the sample size is small ($n < 40$) and σ^2 is unknown, we must make an assumption about the form of the underlying distribution to obtain a valid confidence interval procedure.
- Suppose that the population of interest has a normal distribution with unknown mean and unknown variance σ^2 . Assume that a random sample of size n , say x_1, x_2, \dots, x_n is available, and let

sample mean,

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

and sample variance,

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

- Now we wish to construct a two-sided CI on μ . If the variance is known, it can be shown that

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

has a standard normal distribution, namely, $Z \sim N(0, 1)$.

- When σ^2 is unknown, we replace σ with the sample standard deviation S . The random variable Z now becomes

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

namely, T has a t distribution with $n-1$ degrees of freedom.

Remark: What effect does replacing σ by s on the distribution of random variable T ? Date No. 81

- If n is large, $n > 40$, the answer to this "very little", we can proceed to use the confidence interval based on normal distribution. However, n is usually small in most engineering problems. In this situation a t-distribution must be employed to construct the confidence interval.

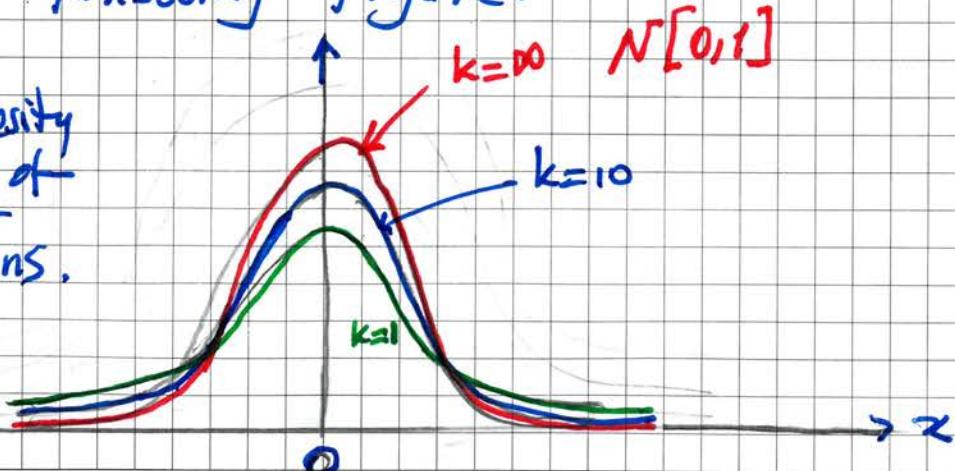
t - Distribution:

Let X_1, X_2, \dots, X_n be a random sample from a normal distribution with unknown μ and unknown variance σ^2 . The random variable

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has a t-distribution with $n-1$ degrees of freedom. Several t distributions are shown in the following figure.

- Probability density functions of several t distributions.



- The mean and variance of the t distribution are

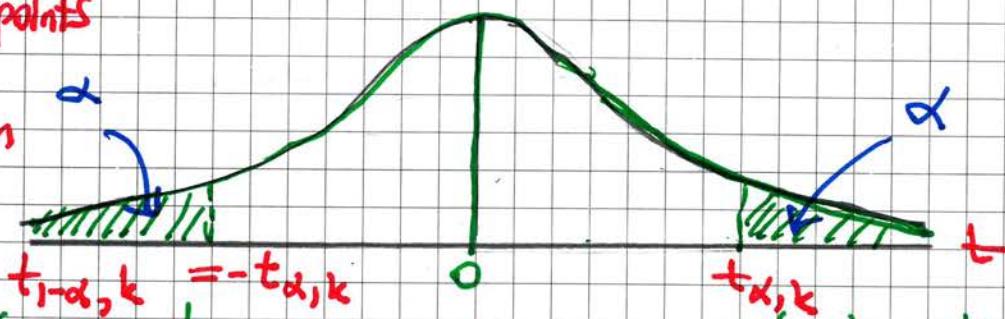
$$\mu_t = 0 \quad \text{and} \quad \sigma_t^2 = \frac{k}{k-2}, \text{ for } k > 2.$$

- t - distribution is similar to the standard normal distribution. However, t distribution has heavier tails than the normal. This means, it has more probability in the tails than the normal distribution.

- As the number of freedom $k \rightarrow \infty$, the limiting form of the t-distribution is the standard normal distribution.

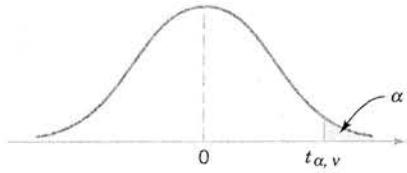
- Let $t_{\alpha, k}$ be the value of the random variable T with k degrees of freedom above which we find an area (or probability) α .

Percentage points
of the
t distribution



Therefore, $t_{\alpha, k}$ is an upper tail 100 α percentage point of the distribution with k degrees of freedom. The percentage points of t - distribution is shown at above.

- In the following table, the α values are the column headings, and the degrees of freedom are listed in the left column.

Table V Percentage Points $t_{\alpha, v}$ of the t-Distribution

$v \setminus \alpha$.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

 v = degrees of freedom.

- For example, the t value with 10 degrees of freedom having area of 0.05 to the right side is $t_{0.05, 10} = 1.812$. Indeed,

$$\Pr\{T_{10} > t_{0.05, 10}\} = \Pr\{T_{10} > 1.812\} = 0.05.$$

- Since the t-distribution is symmetric about zero, we have

$$t_{1-\alpha, n} = -t_{\alpha, n}.$$

Indeed, the value has an area of $1-\alpha$ to the right, and therefore, an area of α to the left. In other words,

$$t_{0.95, n} = -t_{0.05, 10} = -1.812.$$

Note that $t_{\alpha, \infty}$ is the standard normal distribution. You can see in the last row of the t-table, $t_{\alpha, \infty} = Z_\alpha$.

t CONFIDENCE INTERVAL ON μ

The distribution of

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

is with $n-1$ degrees of freedom. Let $t_{\alpha/2, n-1}$ be upper $100(\frac{\alpha}{2})$ percentage. We can write,

$$\Pr \left\{ -t_{\alpha/2, n-1} \leq T \leq t_{\alpha/2, n-1} \right\} = 1-\alpha$$

$$\Pr \left\{ -t_{\alpha/2, n-1} \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{\alpha/2, n-1} \right\} = 1-\alpha,$$

If we rearrange the last equation, we get

$$\Pr \left\{ \bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} \right\} = 1-\alpha.$$

Then we can state that

- If \bar{x} and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance, a $100(1-\alpha)$ percent confidence interval on μ is given

$$\bar{x} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

where $t_{\alpha/2, n-1}$ is the upper $100\frac{\alpha}{2}$ percentage point of the t-distribution with $(n-1)$ degrees of freedom.

Examples [Montgomery and Runger]

The load at specimen failure of a tensile (Gehim) test is given as follows (in megapascals).

$x_1 = 19.8$	10.1	14.9	7.5	15.4	15.4
15.4	18.5	7.9	12.7	11.9	$x_{22} = 11.4$.
11.4	14.1	17.6	16.7	15.8	
19.5	8.8	13.6	11.9	11.4	

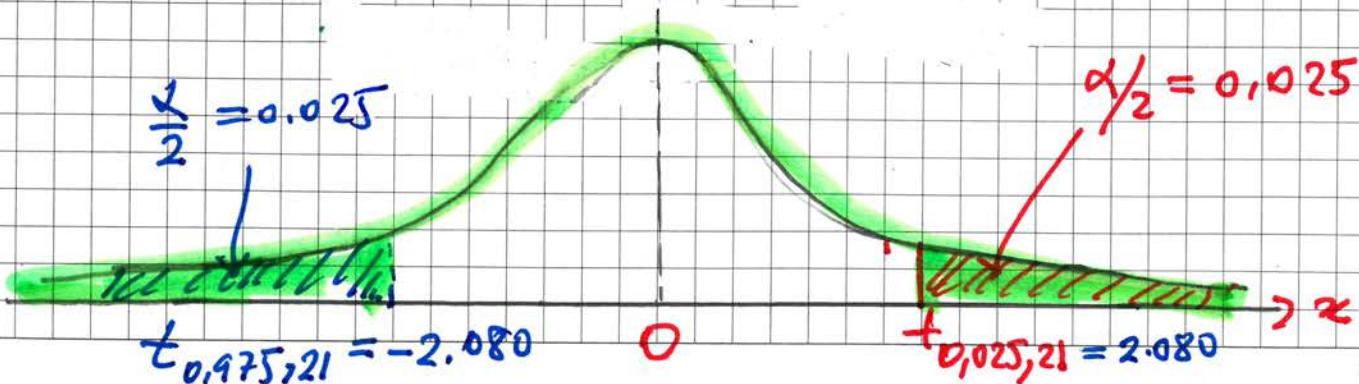
- The sample mean is

$$\bar{x} = \frac{1}{22} \sum_{i=1}^{22} x_i = 13.71.$$

- The sample standard variance is

$$s = \sqrt{\frac{1}{22-1} \sum_{i=1}^{22} (x_i - 13.71)^2} = 3.55.$$

- We want to find a 95% confidence interval on μ . Since $n=22$, we have $n-1=21$ degrees of freedom for t , so, from the t -distribution table, we get $t_{0.025, 21} = 2.080$.



The resulting CI is

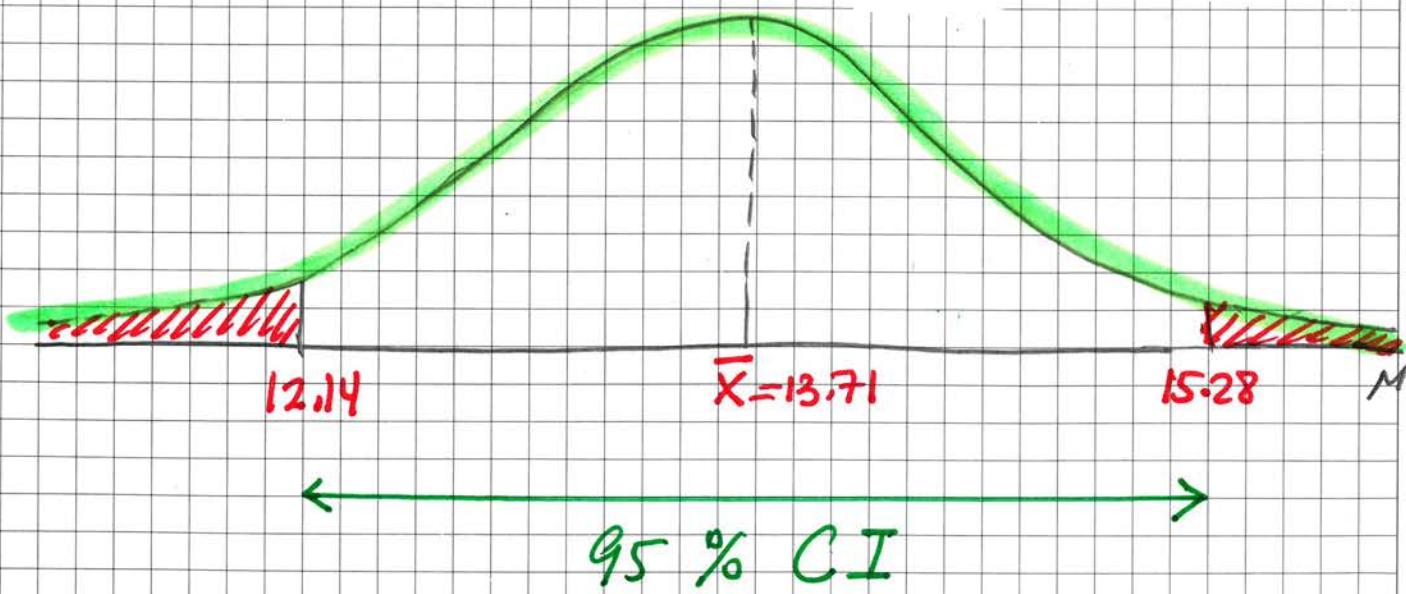
$$\bar{x} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

$$\bar{x} - t_{0.025, 21} \frac{3.55}{\sqrt{22}} \leq \mu \leq \bar{x} + t_{0.025, 21} \frac{3.55}{\sqrt{22}}$$

$$13.71 - 2.080 \frac{3.55}{\sqrt{22}} \leq \mu \leq 13.71 + 2.080 \frac{3.55}{\sqrt{22}}$$

$$13.71 - 1.57 \leq \mu \leq 13.71 + 1.57$$

$$12.14 \leq \mu \leq 15.28$$



CONFIDENCE INTERVAL ON VARIANCE AND STANDARD DEVIATION OF A NORMAL DISTRIBUTION

(chi-square)

Chi-Squared (χ^2) Distribution:

- This distribution arises from the fact that a sum of the squares of n independent standard normal variables has a chi-squared distribution with n degrees of freedom. That is if U_1, U_2, \dots, U_n independent and distributed as $N(0,1)$, the sum

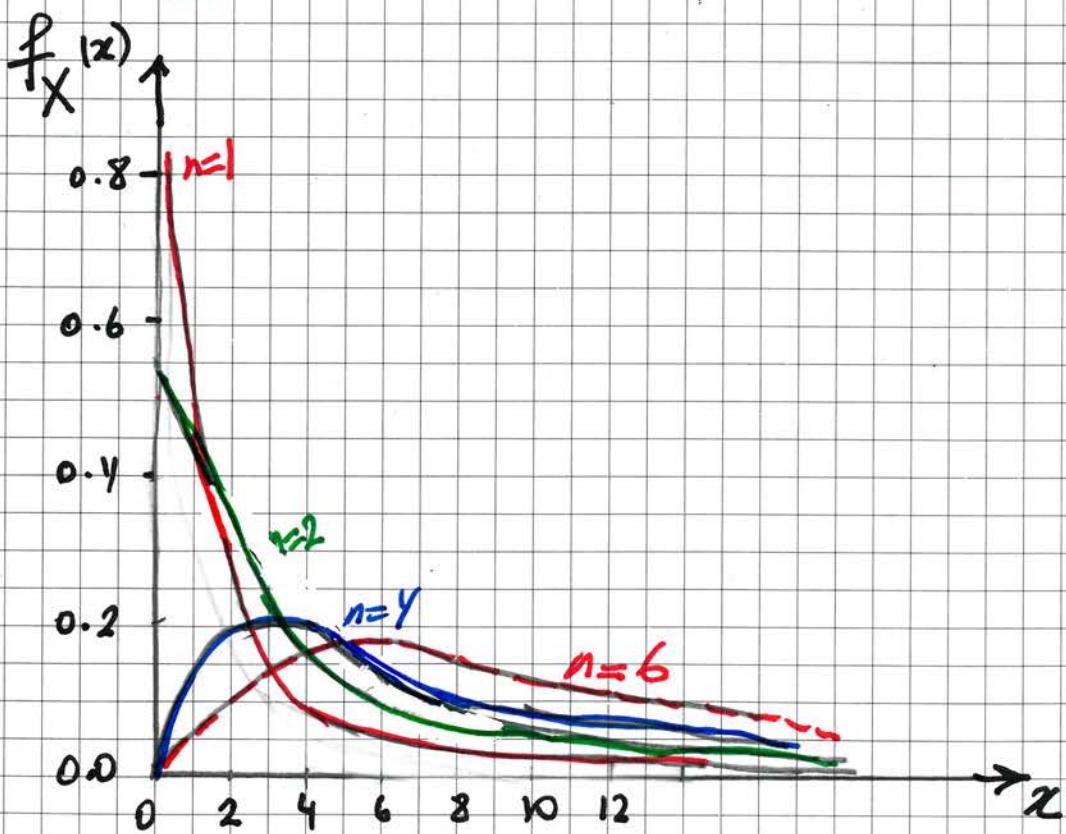
$$X = U_1^2 + U_2^2 + \dots + U_n^2$$

has a χ^2 (chi-square) distribution with n degrees of freedom.

- The pdf of $f_X(x)$ can be plotted for several values of n . It is shown that, as n increases, the shape of $f_X(x)$ becomes more symmetric.

- Since X can be expressed as a sum of identically distributed random variables, we expect that the χ^2 distribution approaches a normal distribution as $n \rightarrow \infty$ on the basis of the central limit theorem.
- The mean and variance of random variable X having a χ^2 distribution are easily obtained as

$$\mu_X = n, \quad \sigma_X^2 = 2n.$$



The χ^2 distribution for $n=1, n=2, n=4$, and $n=6$.

Theorem:

Let X_1, X_2, \dots, X_n be a random sample from a normal distribution with mean μ and variance σ^2 , and let S^2 be the sample variance. Then the random variable

$$\chi^2 = \frac{(n-1)}{2} S^2$$

has a chi-square (χ^2) distribution with $n-1$ degrees of freedom.

- The percentage points of the χ^2 distribution is given in the following table:

Critical values of the Chi-square distribution with k degrees of freedom

k	Probability of exceeding the critical value			k	Probability of exceeding the critical value		
	0.05	0.01	0.001		0.05	0.01	0.001
1	3.841	6.635	10.828	11	19.675	24.725	31.264
2	5.991	9.210	13.816	12	21.026	26.217	32.910
3	7.815	11.345	16.266	13	22.362	27.688	34.528
4	9.488	13.277	18.467	14	23.685	29.141	36.123
5	11.070	15.086	20.515	15	24.996	30.578	37.697
6	12.592	16.812	22.458	16	26.296	32.000	39.252
7	14.067	18.475	24.322	17	27.587	33.409	40.790
8	15.507	20.090	26.125	18	28.869	34.805	42.312
9	16.919	21.666	27.877	19	30.144	36.191	43.820
10	18.307	23.209	29.588	20	31.410	37.566	45.315

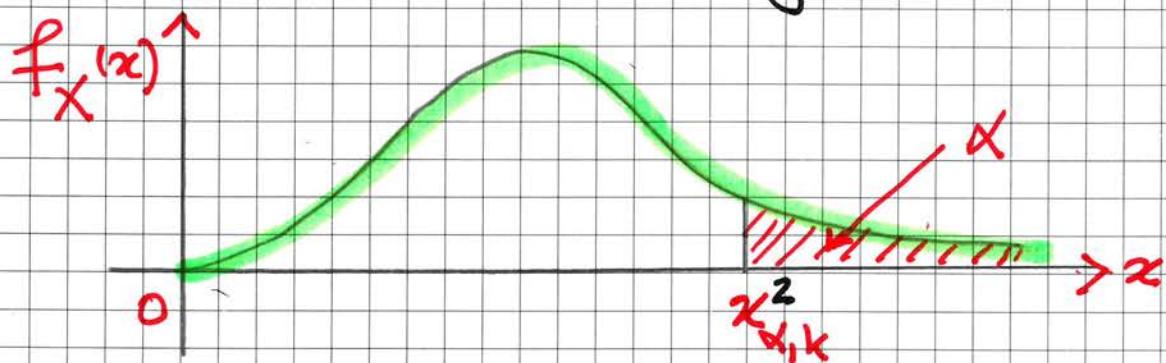
INTRODUCTION TO POPULATION GENETICS, Table D.1

- We define $\chi_{\alpha, k}^2$ as the percentage point of the chi-square random variable with k degrees of freedom such that

the probability of that X^2 exceeds this α value. Namely,

$$\Pr \{ X^2 > \chi_{\alpha, k}^2 \} = \int_{\chi_{\alpha, k}^2}^{\infty} f(u) du = \alpha.$$

- This probability is shown as the shaded area in the following figure:



The percentage point $\chi_{\alpha, k}^2$.

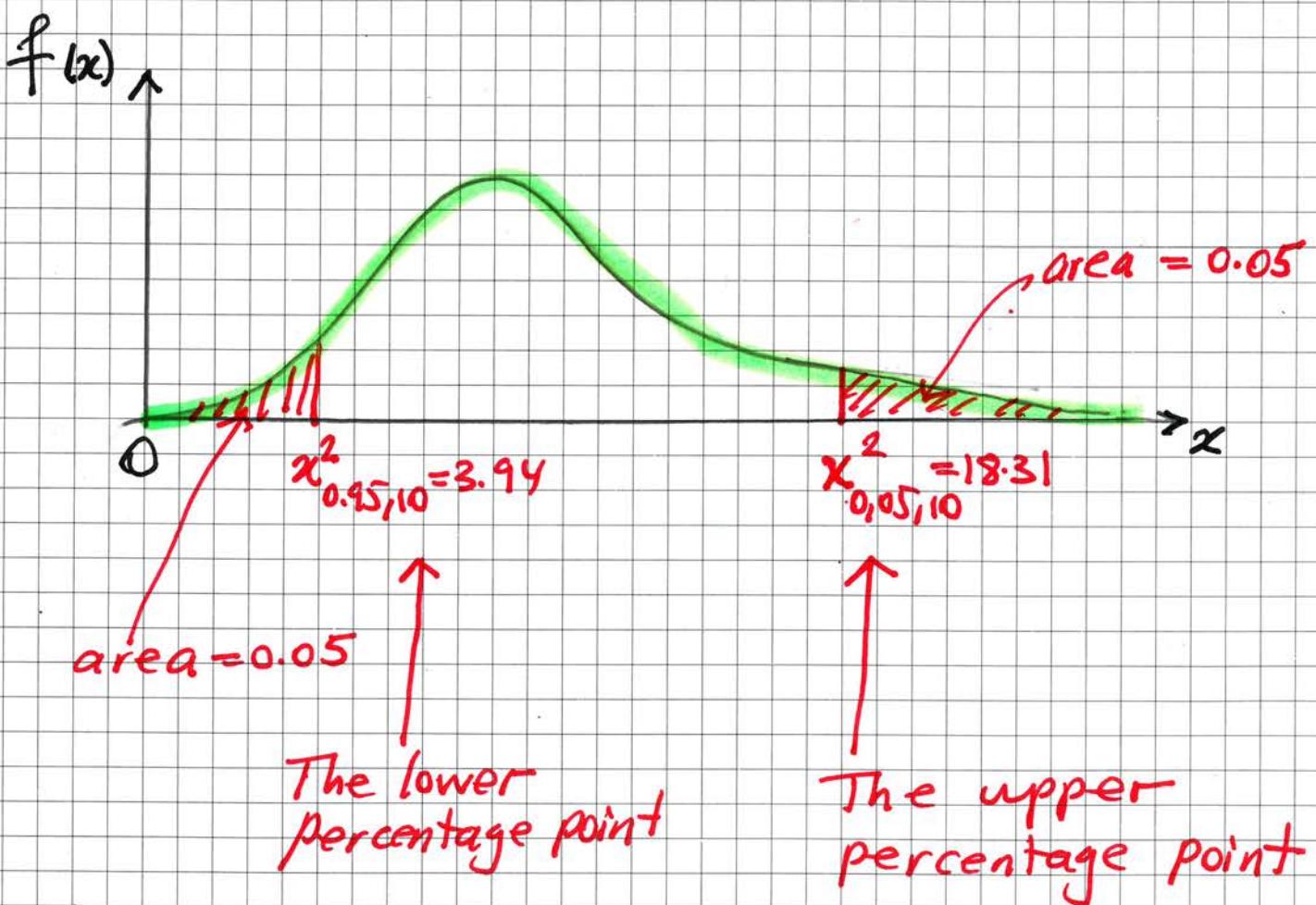
Example:

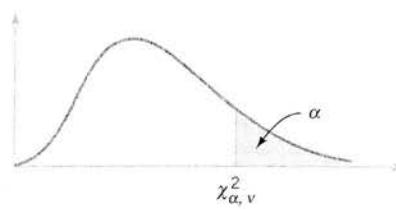
Using the chi-square table, the percentage point $\chi_{\alpha, k}^2$ with 10 degrees of freedom having an area (probability) of 0.05 to the right is $\chi_{0.05, 10}^2 = 18.31$. This value is called as upper 5% point of chi-square with 10 degrees of freedom. Indeed, we can write,

$$\Pr\{X^2 > \chi_{0.05, 10}^2\} = \Pr\{X^2 > 18.31\} = 0.05$$

Conversely, from the table, a lower 5 % point of chi-square with 10 degrees of freedom would be

$$\chi_{0.95, 10}^2 = 3.94.$$



Table IV Percentage Points $\chi^2_{\alpha, v}$ of the Chi-Squared Distribution

$v \setminus \alpha$.995	.990	.975	.950	.900	.500	.100	.050	.025	.010
1	.00+	.00+	.00+	.00+	.02	.45	2.71	3.84	5.02	6.6
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.2
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.3
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.2
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.0
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81

 v = degrees of freedom.

CONFIDENCE INTERVAL on the VARIANCE:

The construction of the $100(1-\alpha)$ CI for σ^2 can be done straightforward.

Since

$$X^2 = \frac{(n-1)S^2}{\sigma^2}$$

is X^2 (chi-square) with $(n-1)$ degrees of freedom, we can write.

$$\Pr \left\{ \chi_{1-\alpha/2, n-1}^2 \leq X^2 \leq \chi_{\alpha/2, n-1}^2 \right\} = 1-\alpha$$

so that

$$\Pr \left\{ \chi_{1-\alpha/2, n-1}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{\alpha/2, n-1}^2 \right\} = 1-\alpha$$

we can rearrange it as

$$\Pr \left\{ \frac{\chi_{\alpha/2, n-1}^2}{\chi_{1-\alpha/2, n-1}^2} \leq \frac{(n-1)S^2}{\sigma^2} \leq \frac{\chi_{1-\alpha/2, n-1}^2}{\chi_{\alpha/2, n-1}^2} \right\} = 1-\alpha.$$

Then we can obtain the following definition of the confidence interval for σ^2 .

If s^2 is the sample variance from a random sample of n observations from a normal distribution with unknown variance σ^2 , then,

a $100(1-\alpha)$ CI on σ^2 is

$$\frac{(n-1)s^2}{\chi_{\alpha/2, n-1}^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_{1-\alpha/2, n-1}^2}$$

where $\chi_{\alpha/2, n-1}^2$ and $\chi_{1-\alpha/2, n-1}^2$ are the upper and lower $100\frac{\alpha}{2}$ percentage points of the chi-square distribution with $n-1$ degrees of freedom, respectively.

Remark:

- We can similarly obtain the one-sided confidence bounds on the variance:
- The $100(1-\alpha)\%$ lower and upper bounds on σ^2 are

$$\frac{(n-1)s^2}{\chi_{\alpha, n-1}^2} \leq \sigma^2 \text{ and } \sigma^2 \leq \frac{(n-1)\sigma^2}{\chi_{1-\alpha, n-1}^2},$$

respectively.

Example: Detergent Filling.

Date.

No. 96

- An automatic filling machine is used to fill bottles with detergent. A random sample of 20 bottles results in a sample variance of fill volume of $s^2 = 0.0153$ (fluid once)².
- If the variance of fill volume is too large, an unacceptable proportion of bottles will be under or overfilled.
- We will assume that the fill volume is approximately normally distributed.
- A 95% upper-confidence interval is found as follows:

$$\sigma^2 \leq \frac{(n-1) s^2}{\chi^2_{0.95, 19}} \quad \text{or} \quad \sigma^2 \leq \frac{(19)(0.0153)}{10.117} = 0.0287 \quad (\text{fluid once})^2.$$

And a confidence interval on the standard deviation is

$$\sigma \leq 0.17 \text{ fluid once.}$$

Therefore, at the 95% level of confidence, the data indicate that the process standard deviation could be as large as 0.17 fluid once.

1 US fluid ounce : 29.57 millilitres.