❖ Probability of occurring Type I error and probability of occurring Type II error are related:

if probability of one type of error __decreases__

then probability of the other type of error __increases__ when the sample size is a constant.

❖ **Probability** of occurring one error can be **reduced** by **increasing sample size** when the probability of occurring the other error is a constant.

➢ Therefore, in hypothesis testing, researcher specifies a __smaller__ value for the

α → **significance level (denoted by α which is the probability of making Type I error)** and use a __larger__ **sample size** so that probability of making Type II error is also small.

1% 5% 10%

➢ Typical values used for **significance level denoted by** α are 0.01, 0.05, and 0.10.

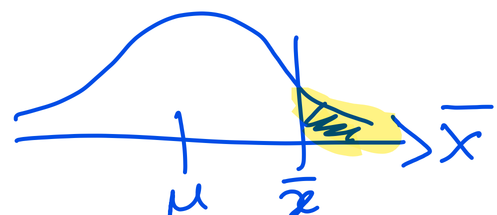
Recall: The typical confidence levels are 99%, 95%, 90%.

For example: If significance level is 5%, then 5 out of 100 hypothesis tests used for the same hypotheses using different samples will make type I error (reject true H_0)

➢ Since **type I error rate can be controlled** by setting significance level, and the **type II error rate is difficult to control**, decide the alternative hypothesis (H_a) so that **type I error is more serious** than type II error.

Ch 10.3: The Use of P-Values for Decision Making in Testing Hypotheses

Steps of a Hypothesis Test (In Practice)



$$\frac{\text{Variable} - \text{Mean}}{\text{Standard Deviation}}$$

1. Write the **null (H_0) and alternative (H_1)** hypotheses.
2. Select a value for the **significance level** (α).
3. Collect data and check the validity of **assumptions**.
4. Compute **test statistic** by standardizing the observed value of the statistic **assuming null (H_0) hypothesis is true.**
5. Then compute the **p-value.**
   P-value is the probability that the statistic is beyond the observed value if the parameter is equal to the claimed value (that is, if H_0 is true).

Decision rule

6. **If p-value is ≤ significance level (α),** then **reject H_0** in favor of H_1 and conclude the data are **significant** (there is **sufficient** evidence for H_1).
   That is, the observed sample did not occur due to just the chance.
   ← p-value > α
   Otherwise, **DO NOT reject H_0** (still considered to be plausible) and conclude the data are **NOT significant** (there is **insufficient** evidence for H_1).

Ch 10.4: Single Sample: Tests Concerning a Single Population Mean μ

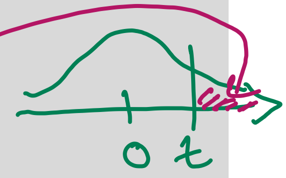
The One-Sample t Test

If the sample can be assumed from a Normal distribution, and population standard deviation (σ) is unknown, then

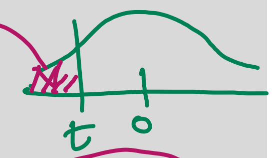
Null hypothesis (H_0) : $\mu = \mu_0$ ← Claimed value

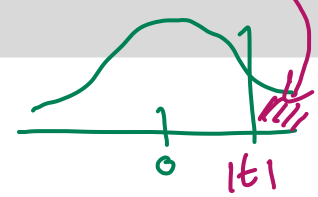
Assumptions: **The sample is random and from a Normal distribution.** and σ is unknown

Test Statistic: $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ ← assuming H_0: $\mu = \mu_0$



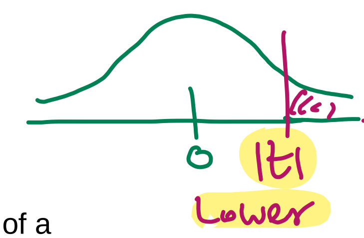
Alternative Hypothesis:	**P-value:**
"Right-tail" H_1: $\mu > \mu_0$	**Area** under the **t** with $(df = n - 1)$ **to the RIGHT of** t
"Left-tail" H_1: $\mu < \mu_0$	**Area** under the **t** with $(df = n - 1)$ **to the LEFT of** t
"Two-tail" H_1: $\mu \neq \mu_0$	**Twice the Area** under the **t** with $(df = n - 1)$ **to the RIGHT of** $\|t\|$





Use the **tcdf function** on the Ti calculator to compute **area** to the **right of** $|t|$ from the t **distribution with** $df = n - 1$

tcdf(|t| , 10^{99} , n−1)



∞
10^{99}
upper

Example 3: In an experiment to investigate the corrosion properties of a particular type of chilled cast iron, a collection of 10 random samples of this chilled cast iron provided corrosion rates with a mean of 2.752 and standard deviation of 0.28. $n = 10$, $\bar{x} = 2.752$, $s = 0.28$

Assume that the corrosion rate of this type of cast iron has a normal distribution.

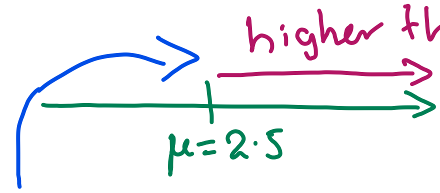
Is there sufficient evidence to conclude that the average corrosion rate of chilled cast iron of this type is larger than 2.5 at a significance level of 0.10?

Claim: $\mu > 2.5$. Here, $\mu_0 = 2.5$ is the claimed value.

Hypotheses:

Null H_0: $\mu = 2.5$

Alternative H_1: $\mu > 2.5$



Since $H_1: \mu > 2.5$, this is a right-tail test.

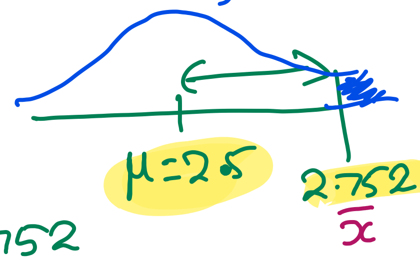
Check Assumptions:

It is given the population has a normal distribution, sample is random and σ is NOT given.
use t-test.

Test Statistic: Assuming $H_0: \mu = 2.5$ is true,

$$t = \frac{\bar{x} - \mu_0}{\left(\frac{S}{\sqrt{n}}\right)} = \frac{2.752 - 2.5}{\left(\frac{0.28}{\sqrt{10}}\right)} = 2.84605$$

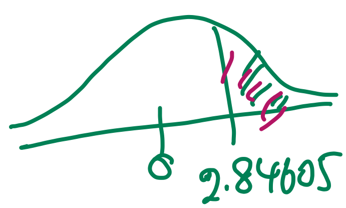
Standard error $\rightarrow \left(\frac{S}{\sqrt{n}}\right)$



$\Rightarrow$ The observed sample mean of 2.752 is 2.84605 standard errors higher than the claimed population mean of 2.5.

P-value:

Since test statistic is positive, compute



$P(T > 2.84605) = 0.0096$

using tcdf(2.84605, 10^{99}, 10-1) $\leftarrow n-1 = 10-1 = 9$

Since this is a right-tail test, p-value $= 0.0096$

$\Rightarrow$ If the null hypothesis was true (that is, $\mu = 2.5$), then there is 0.96% chance that a random sample of 10 such items has a mean of 2.752 or more extrem value.

Decision/Conclusion:

We are given significance level is 0.10
P_value is smaller than significance level.
Reject null hypothesis ($H_0: \mu = 2.5$).
There is enough evidence that population (true) average corrosion rate of chilled cast iron is higher than 2.5. (p-value = 0.0096)

Example 4: The machine that produces metal cylinders is set to make cylinders with a diameter of 50 mm. Isit calibrated correctly?

Regardless of the machine setting, there is always some variation in the cylinders produced.

A random sample of 20 ← n metal cylinders produced from this machine has an average of 49.999mm with a standard deviation of 0.133mm.

Test if the machine is calibrated correctly at a significance level of 0.01 assuming diameter of cylinders produced this machine has a normal distribution.

Let μ be the population mean diameter of the cylinders.

Claim: Machine is calibrated correctly $\Rightarrow \mu = 50$.

Hypotheses:

Null $H_0: \mu = 50$ Machine is calibrated correctly

Alternative $H_1: \mu \neq 50$ Machine is NOT calibrated correctly.

This is a two-tail test because $H_1: \mu \neq 50$ indicates $\mu < 50$ or $\mu > 50$

(Note that if the machine is NOT calibrated properly, μ can be smaller than or higher than 50 and we don't know)

Assumptions:

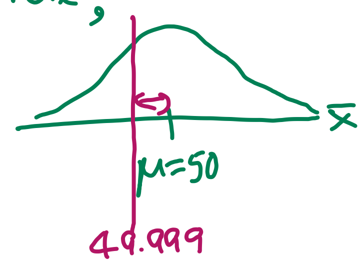
It is given that population has a normal distribution and sample is random and σ is not given.

Use t-test.

Test statistic: Assuming $H_0: \mu = 50$ is true,

$$t = \frac{\bar{x} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{49.999 - 50}{\frac{0.133}{\sqrt{20}}} = \underline{\underline{-0.0336}}$$

standard error → $\left(\frac{s}{\sqrt{n}}\right)$

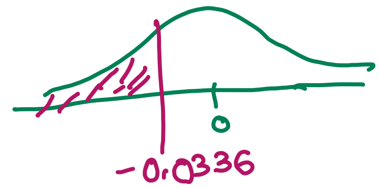


$\Rightarrow$ observed sample mean of 49.999 is 0.0336 standard errors smaller than the claimed population mean of 50.

P-value: Since test statistic is negative, compute

$P(T < -0.0336) = 0.486763$

using tcdf$(-10^{99}, -0.0336, \underbrace{20-1})$ $df = n-1 = 20-1 = 19$



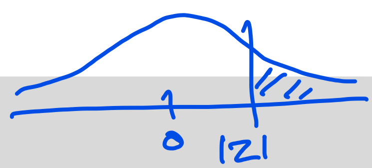
Since this is a two-tail test,

P-value $= 2 * 0.486763 = \underline{0.9735}$

⇒ If the null hypothesis was true (that is, machine was calibrated correctly), then there is 97.35% chance that a random sample of 20 cylinders has a mean of 49.999 or more extreme.

Decision/Conclusion: Given significance level is 0.01 and p-value is higher than the significance level.
Fail to (or Don't) reject the null hypothesis.
There is no enough evidence that the machine is calibrated incorrectly.
⇒ It is plausible the machine is calibrated correctly.

If the sample is selected from a **non-normal distribution**, **n ≥ 30, and population standard deviation (σ) is unknown**, then



normalcdf$(|z|, 10^{99}, 0, 1)$

Assumptions: n ≥ 30

Test Statistic : $z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Approximate P-value is computed from a N(0,1) and make the conclusion as before.

If the sample is selected from an **unknown distribution**, n < 30**, and population standard deviation (σ) is unknown**, then

Plot the data.

If the plots show the sample has approximately normal distribution with no outliers, then use the t-test.

Example 5: A melting point test of 45 samples of a binder use in manufacturing a rocket propellant resultedin average of 154.2^{0}F and a standard deviation of 1.5^{0}F.

Test the claim that true average (μ) melting point of this binder is different from 155^{0}F at significance level of α = 0.05.

claim: $\mu \neq 155$

Null : $H_0: \mu = 155$

Alternative: $H_1: \mu \neq 155$

Two-tail test.