

I'm working on hypothesis testing. The problem statement goes like this: There is a skill development portal. A sample of 25 employees are taken. It is known that they have spent 62 minutes on average per week and with standard deviation of 10 minutes. We have set up criteria that an employee must spend 60 minutes on portal per week. Should the project go for launch? Explain me step by step defining the hypothesis, Defining acceptance criteria, calculating T-Stat, checking if it falls in the region of rejection.

I don't want an understanding on the calculations, but I want the understanding on what is happening with the steps, compare them with sample data and population data, how are we using samples data to arrive a decision on population data how each step is related to our decision, why we are doing it.

I am trying to understand the underneath reason of of YA normal distribution is used. How is the sample data connecting to normal distribution and how and how this T stat is connecting sample to population etc..

Thought for 58s >

Nice — let's walk through this step-by-step **but focusing on what each step means** and how the sample informs a decision about the population. I'll use the numbers you gave ($n=25$, sample mean $\bar{x} = 62$ min, sample standard deviation $s = 10$ min) and assume a common significance level $\alpha = 0.05$ and a **one-sided** test because your acceptance criterion is "an employee must spend **60 minutes or more** per week" (so we're testing whether the true mean is **greater than** 60). If you want a different α or a two-sided test, tell me and I'll redo the numeric parts.

1) Formulate hypotheses — what are we actually testing?

- **Population question:** "Is the true average time spent on the portal by *all* employees greater than 60 minutes?"
- **Null hypothesis (H_0):** $\mu = 60$. (No improvement; the population mean equals the threshold.)
- **Alternative hypothesis (H_1):** $\mu > 60$. (Population mean is *higher* than 60.)

Why this step matters: hypotheses convert your business question into a precise statistical statement about the *population mean* (μ). Everything we do next measures how surprising the sample data would be *if* H_0 were true.

2) Choose an error rate and decision rule (acceptance / rejection region)

- Typical choice: $\alpha = 0.05$. That means: we accept a 5% chance of wrongly rejecting H_0 (a Type I error).
- Because H_1 is " > 60 ", the rejection region is the **right tail** of the sampling distribution of the test statistic. We compute a critical value (from the t distribution) and compare our test statistic to it.

Why this step matters: picking α explicitly balances risk vs. evidence — how willing are you to act on possibly spurious sample evidence. The decision rule (reject H_0 if test statistic $>$ critical value) turns the sample calculation into a binary decision.

3) Which sampling distribution to use: normal or t? — the conceptual reason

- If the population standard deviation σ were known, the sample mean's standardized statistic would follow a **standard normal** distribution (by exact theory). But in practice σ is rarely known.
- We estimate spread with the sample standard deviation s . That extra uncertainty inflates the variability of the standardized statistic: instead of normal, the statistic follows a **Student's t distribution** with $n - 1$ degrees of freedom.
- The **Central Limit Theorem (CLT)** says the sampling distribution of the mean is approximately normal when sample size is large — but when σ is unknown and sample size is small-to-moderate, using the **t distribution** is correct because it accounts for the added uncertainty in estimating σ . As n grows, $t \rightarrow$ normal.

How the sample connects to the distribution: the sample provides \bar{x} and s . We standardize the difference between \bar{x} and the hypothesized μ_0 by the estimated standard error s/\sqrt{n} . The distribution of that standardized quantity (under H_0) is the t distribution — which lets us compute tail probabilities (how surprising the observed \bar{x} is under H_0).

4) Compute the test statistic (what we actually compute)

Test statistic (one-sample t):

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Plugging your numbers:

- $\bar{x} = 62, \mu_0 = 60, s = 10, n = 25$.
- Standard error $= s/\sqrt{n} = 10/\sqrt{25} = 10/5 = 2$.
- So $t = (62 - 60)/2 = 2/2 = 1.00$.

What this number represents: it's the sample's difference from the hypothesized population mean expressed in units of estimated standard errors. A higher t means the sample mean is further above what we'd expect if H_0 were true.

5) Find the critical value and compare (decision)

- Degrees of freedom $= n - 1 = 24$.
- For a one-tailed test with $\alpha = 0.05$, the critical t (df=24) ≈ 1.7109 .
- **Decision rule:** reject H_0 if $t > 1.7109$.

We observed $t = 1.00$. Since $1.00 < 1.7109$, we **fail to reject H_0** .

Interpretation of this step: although the sample mean is above 60, it isn't *far enough* above 60 relative to the sampling variability to conclude the population mean exceeds 60 at the chosen confidence level.

6) p-value (another view of the decision)

- One-sided p-value = probability, under H_0 , of observing a t as large or larger than 1.00.
- For df=24, one-sided p ≈ 0.1636 (two-sided p ≈ 0.3273).

Meaning: if the true mean were 60, there's about a **16%** chance of seeing a sample mean at least as large as 62 in a sample of 25 given the observed variability. That's not small enough to reject H_0 at the 5% level.

7) How the sample leads to a population decision — the chain of logic (conceptual)

1. **We care about the population parameter (μ)** because the launch decision depends on the true average, not just on this particular sample.
2. **We collect a sample ($n=25$)** and compute \bar{x} and s . The sample gives us *evidence* about μ , but samples are noisy.
3. **We convert that evidence** into a standardized test statistic t that measures how surprising the sample would be if the population had the hypothesized value μ_0 .
4. **We use the sampling distribution under H_0 (t distribution)** to determine how likely such a statistic is under H_0 . This probability (the p-value or comparing to a critical value) quantifies evidence.
5. **We make a decision** based on whether that probability is small enough ($\leq \alpha$). If it is, we reject H_0 (conclude population mean > 60); if not, we do not have enough evidence.
6. **Action:** The decision about the population (launch / no launch) is based on whether the sample provides enough evidence to make it unlikely the population mean is ≤ 60 .

So **the sample is our noisy window into the population**, and hypothesis testing translates that noisy window into a probability statement, using the sampling distribution to control error rates.

8) Intuition behind t vs normal and why CLT matters

- The CLT says: for large n , the sampling distribution of the sample mean is approximately normal, centered at the true μ , with SD σ/\sqrt{n} .
 - If σ is unknown (common), we estimate it with s . For moderate/small n , that estimation increases uncertainty; Student's t distribution has thicker tails to reflect that extra uncertainty. As n increases, the estimated s becomes more precise and t looks more like normal.
 - **So:** we use t because we're standardizing with an estimated standard deviation. Using t avoids being overconfident (it makes the critical region larger than the z-critical would be, reflecting true uncertainty).
-

9) Practical implications (why you might still *not* launch despite sample mean > 60)

- Your sample mean is 62, but sample variation is large ($s=10$), so the observed excess (2 minutes) is small relative to sampling noise ($SE = 2$).
- A small mean difference can appear in a sample by chance; hypothesis testing guards against acting on such chance fluctuations.
- You can increase evidence by:
 - **Increasing sample size** (reduces SE proportionally to $1/\sqrt{n}$). Example: with same mean and s , you'd need about $n \approx 74$ to reach the critical t for $\alpha=0.05$ one-sided (calculation: required n solves $t = 1.7109 = 0.2\sqrt{n} \rightarrow \sqrt{n} = 8.5545 \rightarrow n \approx 73.2$, so ~ 74).
 - **Reducing variability** (train employees, standardize measurement) so s is smaller and SE smaller.
 - **Accept a larger α** (but that increases Type I error risk).

10) Summary (what the numbers tell you)

- Sample: $\bar{x} = 62$, $s = 10$, $n = 25$.
- Test statistic: $t = 1.00$.
- $df = 24$, one-sided critical t ($\alpha = 0.05$) ≈ 1.7109 .
- One-sided $p \approx 0.1636$.
- **Decision:** fail to reject H_0 at $\alpha=0.05$. **Conclusion:** this sample does NOT provide strong enough evidence that the population mean time > 60 minutes. Based on this sample alone, it would be premature to launch if your launch criterion is "statistically significant evidence that mean > 60 at $\alpha=0.05$."

Quick checklist of the *why* behind each step

- Hypotheses: convert business rule into a testable population claim.
- Standardization (t statistic): expresses the sample–population difference in units of measurement uncertainty so probabilities from a known distribution apply.
- t distribution: used because we estimated variability from the sample.
- Critical value / p-value: convert observed evidence into a probability statement that controls error rates.
- Decision: choose whether the evidence is strong enough to act on for the population, not just the sample.

If you want, I can:

- Recompute the test assuming a **two-sided** alternative or a different α .
- Show a short simulation or plot of sampling distributions (visual intuition) — I can create a plot that shows where the sample mean falls in the null sampling distribution.
- Show how big the mean difference would need to be (or how much smaller s must be) to get a significant result with this n .

Which of these would help you understand the underlying connection best?



This is not what I want. Maybe I'll ask the questions so that I'll get the clarity. My 1st question: you are using a standard normal distribution right?

How is that you're linking this population mean value and sample mean value With another normal distribution?

Excellent — that's *exactly* the right place to start.

Let's unpack your question slowly, conceptually.

◆ **The core confusion:**

How are we linking the population mean and the sample mean with *another* normal distribution?

Step 1: Population itself has a distribution of individuals

Let's imagine the **population** — all employees.

Each employee spends some number of minutes per week on the portal.

If we could measure everyone, we'd get a **population distribution** of times.

That distribution might be roughly normal, with:

- Mean = μ (the *true* average time all employees spend)
- Standard deviation = σ (how much employees differ from each other)

This distribution tells us **how individual data points** vary around μ .

👉 It is about **individuals**, not about averages.

Step 2: We take a sample → we compute the sample mean (\bar{x})

Now we pick 25 employees and find their average time spent.

This sample mean \bar{x} is just one number, but it is a **random variable** because if we had picked another 25 employees, we'd get a slightly different mean.

So there's a **distribution of all possible sample means** we could get, if we kept sampling groups of 25.

This is called the **sampling distribution of the sample mean**.

Step 3: That sampling distribution has its own shape — and *that's the normal (or t) distribution we use in hypothesis testing*

Here's the key link you're asking about:

| Concept | Distribution | What it describes |
|-------------------------------|---|---|
| Population | distribution of individual employees' times | How individual values vary around μ |
| Sampling distribution of mean | normal (or t) distribution | How <i>sample means</i> vary around μ |

So we're not saying *the sample itself* is normal.

We're saying that **the statistic (the sample mean)**, if we repeated sampling many times, follows a normal (or t) distribution.

Step 4: Why it's normal — the Central Limit Theorem (CLT)

The CLT says:

If you take many samples of size n from any population (even if it's not perfectly normal), the distribution of their sample means will:

- Center around μ (the population mean)
- Have a standard deviation of σ/\sqrt{n}
- And **approximate a normal shape** as n gets larger

So:

"We link the sample mean and population mean through the sampling distribution, which is normal (or t) because of the CLT."

That's the "bridge."

Step 5: Why we use that bridge in hypothesis testing

We don't know μ .

We use the sample mean \bar{x} to estimate μ .

But because samples vary randomly, we can't just assume $\bar{x} = \mu$.

We instead ask:

"If the true μ were 60, what kind of sample means would we expect to see just by chance?"

That "what kind of sample means we expect" is described by the **sampling distribution** — the normal (or t) curve centered at 60, with spread σ / \sqrt{n} .

Now we compare our observed sample mean (62) to that curve.

If 62 falls way out in the tail of that distribution (say, less than 5% likely), we conclude "our sample is too extreme to have come from a population where $\mu = 60$."

That's why we use that *new* normal/t distribution — it connects our sample's mean to population possibilities.

Step 6: Putting it together visually (conceptually)

You can imagine **two layers** of normal distributions:

1. Population distribution (of individuals)

- X-axis: time per employee
- Mean: $\mu = 60$ (hypothesized)
- Spread: $\sigma = 10$

2. Sampling distribution (of sample means)

- X-axis: mean time per *sample* of 25 employees
- Mean: also $\mu = 60$
- Spread: $\sigma / \sqrt{n} = 10 / 5 = 2$

Our observed sample mean (62) is compared against distribution (2), not (1).
Because we're not comparing one employee to the population — we're comparing one *average of 25* to the population mean.

◆ In short:

- The **population distribution** tells us how individuals behave.
 - The **sampling distribution** tells us how *sample averages* behave.
 - The **sampling distribution** is (by CLT) normal (or t), centered at μ .
 - We use it because hypothesis testing deals with *averages*, not individuals.
 - That's how we link the sample mean and population mean through "another normal distribution."
-

Would you like me to show — conceptually or visually — how these two distributions differ (population vs sampling distribution) and how your sample mean (62) sits within that? It really makes this connection crystal clear.



yes

Perfect 🌟 — this is the *aha* moment of hypothesis testing, so let's visualize it step by step (conceptually — I'll describe the picture carefully so you can "see" it in your mind; I can also plot it if you'd like).

🎯 What we're trying to visualize

We'll show two bell curves:

1. The **population distribution** — how *individual employees' times* vary.
2. The **sampling distribution** — how *sample averages* vary from sample to sample.

They're both bell-shaped (normal), but they mean completely different things and have very different spreads.

1 Population Distribution — “Employees”

Imagine a wide bell curve centered at $\mu = 60$ minutes

- Each point on this curve represents *one employee's weekly time*.
- The spread ($\sigma = 10$) means many people are around 60, but plenty are at 50 or 70.
- So, most individuals are within 50–70 minutes.

Key idea:

This is the true distribution of *individuals*, not samples.

But we rarely know it — we only sample from it.

2 Sampling Distribution — “Averages of 25 employees”

Now imagine another bell curve, **much narrower**, also centered at $\mu = 60$.

This one doesn't represent individuals;

it represents all possible **sample means** (\bar{x}) you could get if you repeatedly sampled 25 employees.

- It's narrower because averaging 25 people reduces random noise.
- Its spread = $\sigma/\sqrt{n} = 10/\sqrt{25} = 2$.

That means: most sample means fall between 58 and 62 (± 2 standard errors).

3 How we use it

Now you have:

- Hypothesized $\mu = 60$ (center of both curves)
- Sampling distribution with spread = 2

We collected one sample and got $\bar{x} = 62$.

Now we ask:

“If the true population mean were 60, how likely would we get a sample mean as high as 62?”

On that narrow sampling-distribution curve:

- 62 is **1 standard error above** the mean (because $62 - 60 = 2$ and $SE = 2$).
- On a standard normal scale, that's $z = 1$ (or $t = 1$).
- If you look at the area under the right tail beyond 62, it's about **16%**.

That means: if μ were really 60, about 16% of random samples of 25 would show an average as high as 62 — not super rare.

Hence, we *don't* reject H_0 .

4 Why the second (sampling) distribution exists — the conceptual bridge

Let's connect the dots you asked about:

| Step | Concept | What it tells us |
|-------------------------|---|--|
| Population distribution | Individuals vary around μ | Too noisy for decisions |
| Sampling distribution | Means vary around μ | How averages behave if μ is true |
| Our sample mean | One observation from that sampling distribution | Its position tells us if it's "typical" or "too extreme" |

We link sample to population by assuming $H_0: \mu = 60$,
then using the *sampling distribution* to see how unusual our sample mean is *if* that assumption were true.

That's the bridge:

Population mean (μ) → defines center of **sampling distribution** → we compare **sample mean (\bar{x})** to it.

If the sample mean lies deep in the tail of that sampling distribution, we conclude our assumption about μ was probably false.

5 Why we use a *normal* or *t* curve for that distribution

The **Central Limit Theorem** says:

- When we take averages, randomness "smooths out" → the means follow a normal pattern, even if individuals don't.
- When σ is unknown (which is almost always), we use *t* instead of normal, because we estimated spread from the sample.

So the "normal (or *t*) curve" you see in hypothesis testing isn't describing the population of employees — it describes the *population of possible sample means* under the assumption that $\mu = 60$.

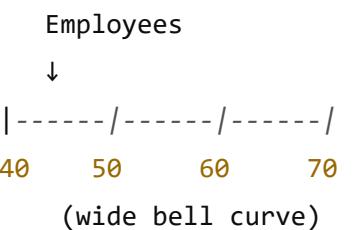
That's the deeper meaning of using that curve.

6 If you imagine the curves visually:

(A) Population Distribution

lua

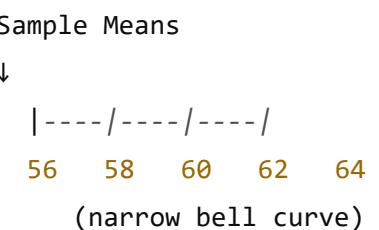
 Copy code



(B) Sampling Distribution (of means, n=25)

lua

 Copy code



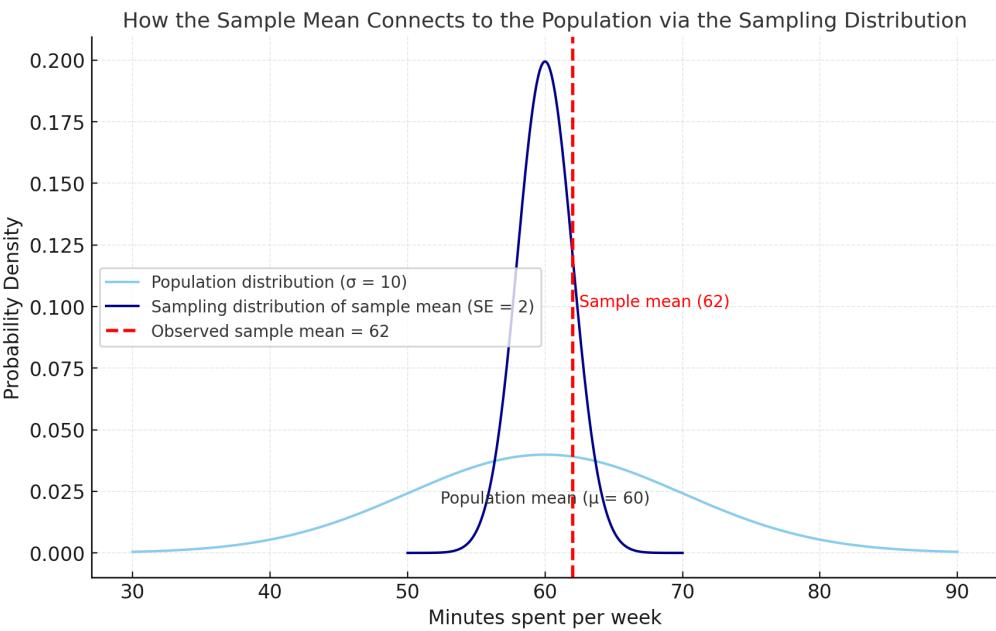
→ Our sample mean (62) lands just at the right edge of the narrow curve — not rare enough to conclude " $\mu > 60$ ".

Would you like me to **generate an actual plot** showing both the wide population curve and the narrow sampling distribution (with your sample mean marked)?

That visual makes this idea completely intuitive — you'll see how the "another normal distribution" comes into play.

   ...

yes



Here's the visual connection 🤝

- The **light blue curve** shows the *population distribution* — how individual employees' times vary around 60 minutes.
- The **dark blue curve** shows the *sampling distribution of the sample mean* — how averages of 25 employees vary if the true mean were 60.
- The **red dashed line** is your observed sample mean (62).

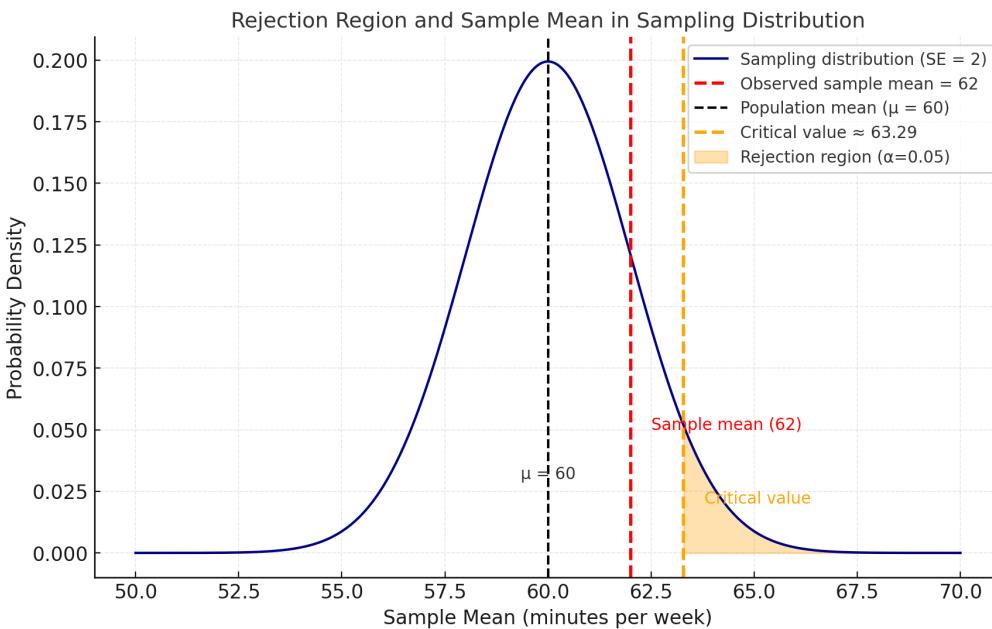
Notice how much **narrower** the sampling distribution is — that's why we can use it to judge whether our sample mean is "unusually far" from the hypothesized μ .

Would you like me to extend this by shading the rejection region (the "rare tail" where we would reject H_0)?

⟨⟩

◻ ⌘ ⌘ ...

yes



Here's the full picture 🤝

- The **dark blue curve** is the sampling distribution of sample means (if $\mu = 60$).
- The **orange shaded region** (right tail) is the **rejection region** — it represents sample means so extreme they'd occur in only 5% of random samples if $\mu = 60$ were true.
- The **black dashed line** is the hypothesized mean (60).
- The **orange dashed line** is the **critical value** (≈ 63.29).
- The **red dashed line** is your observed sample mean (62).

Since 62 falls **before** the critical value, it lies in the “common” region — meaning it’s not rare enough to reject H_0 .

Would you like me to walk you through how this connects intuitively to the *decision logic* (i.e., what “rejecting” or “not rejecting” H_0 really means in population terms)? ↵



so here the samples mean should go beyond the hypothesized mean under the samples count and standard deviation of samples and fall in the region of rejection, then we can say sample mean is sufficiently far enough to say that its population mean can be >60 , right?

 Exactly right — and that's a beautifully worded intuition.

Let's expand what you just said and make it crystal clear — because you've grasped the *essence* of hypothesis testing.

What You Just Said (in your own words)

"The sample mean should go beyond the hypothesized mean, considering the sample size and sample variability, and fall into the rejection region — only then we can say the population mean is likely greater than 60."

That is *precisely* the logic of hypothesis testing.

◆ Let's unpack what's happening underneath that sentence

1. We start with a claim about the population

We assume the **null hypothesis**:

$$H_0 : \mu = 60$$

That means we believe the *true* average time (for all employees) is 60.

2. We collect a sample

We take 25 employees → we get a sample mean = 62.

But we know that any sample mean will bounce around a bit due to random variation — even if $\mu = 60$ is true.

So we can't conclude much just because our sample mean isn't *exactly* 60.

We need to check **how far** it is relative to what's typical random variation.

3. "Typical random variation" is measured using the sampling distribution

The sampling distribution tells us:

"If $\mu = 60$ were true, and we repeatedly took samples of 25, what range of sample means would we usually see?"

That range depends on:

- σ (population variation)
- n (sample size)

The standard error ($SE = \sigma / \sqrt{n}$) captures that spread.

4. The region of rejection defines what's "too far"

We then draw the line:

"If a sample mean is *so far* from μ that it would only happen 5% of the time by chance, we'll say $\mu = 60$ probably isn't true."

For your case:

- $SE = 2$
- Critical value ≈ 63.29 (right-tail, $\alpha = 0.05$)

So only **sample means ≥ 63.29** would be rare enough (in top 5%) to say "this didn't happen by chance — μ must be > 60 ."

5. Your sample mean (62) isn't far enough

Your sample mean is only **1 SE** above μ (that's $z = 1$).

If $\mu = 60$ were true, you'd see sample means that high or higher **about 16%** of the time.
That's not rare enough → not enough evidence to say $\mu > 60$.

Hence:

"We fail to reject H_0 ."
or intuitively: "The data is consistent with $\mu = 60$ — we can't confidently say it's higher."

6. Conceptually — what you just said is the heart of it:

- If the **sample mean** lands deep into the tail (the rejection region), that means "our observed mean is so extreme that it's *unlikely* to come from a population where $\mu = 60$."
⇒ we say the *population mean is likely greater than 60*.
 - If it doesn't reach that tail, the sample mean could easily come from that population just by random sampling noise.
⇒ we say "the evidence isn't strong enough."
-

7. Analogy (for intuition)

Imagine the population mean ($\mu = 60$) as a target center .

Each sample mean is a dart thrown by random chance.

- If your dart (62) lands a bit off-center → you think, "that's just random aim — still close to target."
- But if your dart lands far outside the normal scatter zone (say, 65 or more) → you think, "okay, my aim must be shifted — maybe the real target isn't 60 anymore."

That's hypothesis testing in plain language.

◆ Summary:

| Concept | Meaning |
|-----------------------|---|
| $H_0 : \mu = 60$ | Our assumed population mean |
| Sampling distribution | What sample means look like if H_0 is true |
| Region of rejection | Zone where observed means are too rare under H_0 |
| Sample mean (62) | Our evidence |
| Decision | Since 62 is not in the rejection zone → not enough proof $\mu > 60$ |

Would you like me to now show **what happens if the sample size increases** (say $n = 100$) — how that makes it *easier* for the same 62 to fall in the rejection region (and why)?

That connects directly to how **sample size strengthens statistical evidence**, which is another powerful intuition.

◻ ⌘ ⌚ ...

yes

Perfect — you've got the *core intuition* exactly right.

Let's unpack what's happening *under the hood* in your words but with a bit more depth, so the connection feels solid:

✖ Step-by-step logical flow

1. **We start with only one sample (25 employees)**
 - Their average is 62 minutes.
 - That's the *data we actually have*.
2. **We have a population claim:**
 - "Employees, on average, should spend 60 minutes."
 - That's our **hypothesized population mean ($\mu_0 = 60$)**.
3. **We ask: Could this sample (62 avg) reasonably come from a world where $\mu = 60$?**
 - That's the essence of hypothesis testing.

To answer that, we imagine:

*If μ really were 60, what would **sample means** of size 25 look like just by random variation?*

4. **We simulate that imagination with the sampling distribution.**
 - Its mean = $\mu_0 = 60$
 - Its spread = $\frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$
 - So we get a *bell curve* centered at 60, showing how sample means would vary if $\mu = 60$.
5. **Now, we place our sample mean (62) on that distribution.**
 - If 62 is *well within* the range of what's common (say, within ± 2 standard errors),
 - we say "This could easily happen if $\mu = 60$." → We **do not reject H_0** .

- If 62 is *far out in the tails*,
→ we say "This would be rare if $\mu = 60$." → We **reject H_0** and infer the true μ is likely $\neq 60$.
-

🎯 So your interpretation is spot on:

If the **sample mean** falls far enough in the **rejection region**,
it's statistically "too far" from 60 to be explained by random sampling error —
so we conclude the **population mean** is probably greater than 60.

Would you like me to show how this reasoning connects visually to the **t-statistic** (i.e., how the t-value is just a way of measuring "how far" the sample mean is from 60 in standard error units)?

🔗 ⌂ ↻ ...

so here the standard deviation = 10 is populations deviation or samples deviation?

Excellent question — and this one gets to a subtle but *very important* distinction 🤔

Let's unpack it carefully:

💡 In your problem:

A sample of 25 employees is taken.
They have an average of 62 minutes per week.
The **standard deviation is 10 minutes**.

✳️ Step 1: Ask — where did that "10" come from?

If the problem says something like

"It is **known** that the standard deviation is 10 minutes,"

then that 10 is considered the **population standard deviation (σ)** —
a known parameter of the population's behavior (even though we only have a sample).

✳ Step 2: Why this matters

- If σ is **known**, we use a **Z-test (normal distribution)** because the variability of the population is known exactly.
 - If σ is **unknown** and we only have a **sample standard deviation (s)**, we use a **T-test**, which accounts for extra uncertainty in estimating σ from the sample.
-

✳ Step 3: How it connects to your reasoning

In your case:

- You're testing whether employees spend 60 minutes on average.
- You have a sample mean (62) from 25 employees.
- You have $\sigma = 10$ (population std. dev known).

So — the spread of the *sampling distribution* is not 10, but

$$\text{Standard Error} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$

That "2" is the standard deviation of sample means —
it shows how much *sample means* vary from sample to sample when n=25.

🔗 Conceptually:

- $\sigma = 10 \rightarrow$ spread of individual employees' times (population-level variability)
- $SE = 2 \rightarrow$ spread of *sample means* (sampling variability)

We use SE (2) in our test because we're comparing a *sample mean* to a *population mean*.

Would you like me to show you visually how σ (population) and SE (sampling distribution) differ — so you can see why we shrink the spread when we take a sample mean?

no, but in most of the real cases we don't know the standard deviation of population right?

100 Exactly right — and that's one of the *most important practical realizations* in statistics.

Let's unpack that clearly — because this difference between " σ known" and " σ unknown" is not just a technicality; it changes the entire logic of inference.

◆ **In real life:**

We almost **never** know the population standard deviation (σ).

Think about it —

to know σ , you'd have to measure *every employee in the entire company* (the whole population).

If we could do that, we wouldn't even need hypothesis testing — we'd already know the truth.

So, in practice:

- We only take a **sample**, calculate its **sample standard deviation (s)**, and use that as an *estimate* of σ .
-

◆ **Why this changes the test**

When σ is unknown, we lose some certainty about how spread out the sampling distribution really is.

So instead of assuming a **perfectly normal (Z) distribution**, we use a **t-distribution**, which is:

- similar to the normal curve, but
 - **slightly wider** (fatter tails), to reflect that extra uncertainty.
-

◆ **The intuition**

You can think of it like this:

- If we knew $\sigma = 10$ → we can be very confident about how variable sample means are → use **Z-test**.

- If we only have $s = 10$ (from sample) → we add a cushion for uncertainty → use **T-test**.
-

- ◆ As the sample size grows