Chapter 6

Statistical inference
for the population mean

What we'll cover...

- The **central limit theorem**
- **Confidence intervals**
- One sample **hypothesis tests** for the mean
- Two sample **hypothesis tests** for the mean

6.1 Introduction

What does our sample tell us about the population?

- We can rarely observe the entire population, so the **population mean** and **population variance** are hardly ever known *exactly*;
- These unknown quantities are called **parameters**;
- We use Greek letters to denote them – μ for the mean, and σ^2 for the variance (and so σ for the standard deviation);
- We hope that the **sample mean** ($\bar{x}$) will be quite close to the true mean (μ);
- **But how do we know if it is?**

Example: *The Vintage Clothing Co.*

The *Vintage Clothing Co.* are a large retailer of bespoke and retro clothing.

They have 1,000 branches across the U.K., and all of their branches are open on Sundays.

However, they are considering whether or not it is worthwhile staying open on Sundays.

The table overleaf shows the number of transactions at each of their shops on Sunday 24th February 2013.

Example: *The Vintage Clothing Co.*

282	258	399	271	343	285	247	513	171	123	168	327	430	240	410	341
290	263	446	185	330	111	243	376	139	351	311	389	546	321	393	487
264	320	217	257	349	640	97	298	393	454	363	354	360	326	199	502
293	407	362	270	344	263	290	263	50	253	345	581	229	264	304	394
499	276	412	323	310	177	248	178	409	275	278	307	495	515	232	432
339	404	371	262	336	218	274	483	211	245	316	381	432	233	223	447
202	133	356	408	224	379	197	278	235	509	171	232	429	315	326	602
242	389	219	206	393	437	306	152	294	271	230	398	346	344	379	347
305	174	291	261	214	532	335	63	100	357	190	347	208	420	322	463
389	236	445	378	255	301	308	150	289	453	464	273	211	450	222	250
320	420	357	160	372	99	316	218	248	322	145	399	433	393	403	361
261	279	369	342	168	322	304	254	99	503	303	212	105	166	257	422
346	370	235	355	65	340	420	338	568	644	164	288	319	159	324	208
268	340	305	361	319	519	293	380	286	431	402	329	363	330	612	248
446	588	304	454	164	240	293	478	540	339	245	257	222	471	469	273
277	216	555	401	380	338	212	476	77	363	140	451	329	66	217	461
522	111	119	316	116	471	142	336	277	101	518	264	226	256	539	324
333	332	404	362	202	204	341	80	333	267	439	136	343	389	244	370
372	595	314	182	470	192	555	374	368	192	225	321	435	403	316	312
192	63	407	125	253	89	70	186	491	342	122	367	106	334	161	177
454	122	286	39	361	262	316	272	285	201	191	162	229	334	278	231
644	297	398	118	246	148	478	167	337	344	395	334	255	401	504	304
192	507	41	457	405	306	282	446	195	512	252	510	557	191	321	404
377	240	441	308	346	265	375	332	580	130	353	426	95	588	332	109
263	529	172	529	315	257	481	260	297	382	438	64	226	185	369	275
190	340	337	224	363	212	371	229	175	388	332	315	389	452	266	393
219	400	378	241	616	551	359	489	314	450	645	224	320	405	182	251
240	471	293	240	184	296	617	565	206	147	169	401	140	462	389	310
323	351	187	544	387	425	353	175	378	484	205	295	413	189	559	251
213	574	579	325	246	206	419	306	471	264	270	300	278	131	561	328
281	403	256	348	183	161	444	482	338	268	313	252	179	414	444	266
203	269	450	322	459	183	212	242	144	406	401	174	605	270	487	494

Example: *The Vintage Clothing Co.*

Suppose the marketing department of *The Vintage Clothing Co.* are interested in the **average number of transactions** across all their stores on Sunday 24th February 2013.

Can we work this out *exactly*?

The answer is "yes" — we have data from every single branch!

Actually, the table of results shows we have taken a **census** – every single branch has been asked to provide us with data.

So in this case, it is possible to work out the **population mean** μ:

$$\mu = \frac{282 + 258 + 399 + 271 + \ldots + 426 + 477}{1000} = 320 \text{ transactions.}$$

Example: *The Vintage Clothing Co.*

Now, let's suppose the company don't have the **time**/**resources** to take a census.

In fact, a week before the 24th February, just **five stores** are selected at random — we work out the mean number of transactions using the data from these five stores only.

Let's suppose the top left–hand block of the table is for stores 1–100, the next block is for stores 101–200, etc.

We put the numbers 1–1000 into a bag and draw, without replacement, **5 numbers at random**.

Example: *The Vintage Clothing Co.*

Store	No. of transactions, X
637	$x_1 = 374$
327	$x_2 = 452$
849	$x_3 = 271$
666	$x_4 = 419$
680	$x_5 = 643$

Let's suppose this is the only information we have.

It is no longer possible to work out the true population mean, as we don't have information from every single shop; we can now only work out the **sample mean** $\bar{x}$:

$$\bar{x} = \frac{374 + 452 + 271 + 419 + 643}{5} = 431.8 \approx 432 \text{ transactions.}$$

Example: *The Vintage Clothing Co.*

Obviously, the marketing team are not just interested in what goes on in these five shops.

However, this is the only information they have, and so they use this information to draw conclusions about all 1,000 shops as a whole.

This is known as the process of **statistical inference** – we are trying to *infer* things about the population, based on the limited information in our sample.

Hopefully, provided we don't have a **biased sample**, $\bar{x}$ will do a good job at estimating μ. Has it done a good job here?

Example: *The Vintage Clothing Co.*

The true **population mean** is $\mu = 320$.

Our **sample mean** is $\bar{x} \approx 432$.

We have **over–estimated** the population mean by quite a bit – our estimate is 112 too high!

Perhaps we have a **biased sample** – we seem to have selected a number of stores with a higher–than–average number of transactions – especially store 680!

Example: *The Vintage Clothing Co.*

This begs the question:

"*How accurate can our sample mean be in estimating the population mean?*"

Let's take another random sample by drawing another five numbers from the bag:

Store	No. of transactions, X
558	$x_1 = 253$
428	$x_2 = 446$
903	$x_3 = 251$
364	$x_4 = 256$
14	$x_5 = 185$

This gives

$$\bar{x} = \frac{253 + 446 + 251 + 256 + 185}{5} = 278.2 \approx 278 \text{ transactions.}$$

Example: *The Vintage Clothing Co.*

This sample mean is much closer to the population mean, but still not *very* close.

Also, it is quite different from the mean of the previous sample.

You could repeat this procedure yourself (filling in the table on page 142) to select three more random samples of size 5, and calculate the sample means.

In fact, consider this **homework** to do before the practicals next week!

Your random sample 1

$u =$	$store =$	$x_1 =$
$u =$	$store =$	$x_2 =$
$u =$	$store =$	$x_3 =$
$u =$	$store =$	$x_4 =$
$u =$	$store =$	$x_5 =$
		$\bar{x} =$

How close are these sample means to the correct population value $\mu = 320$ transactions?

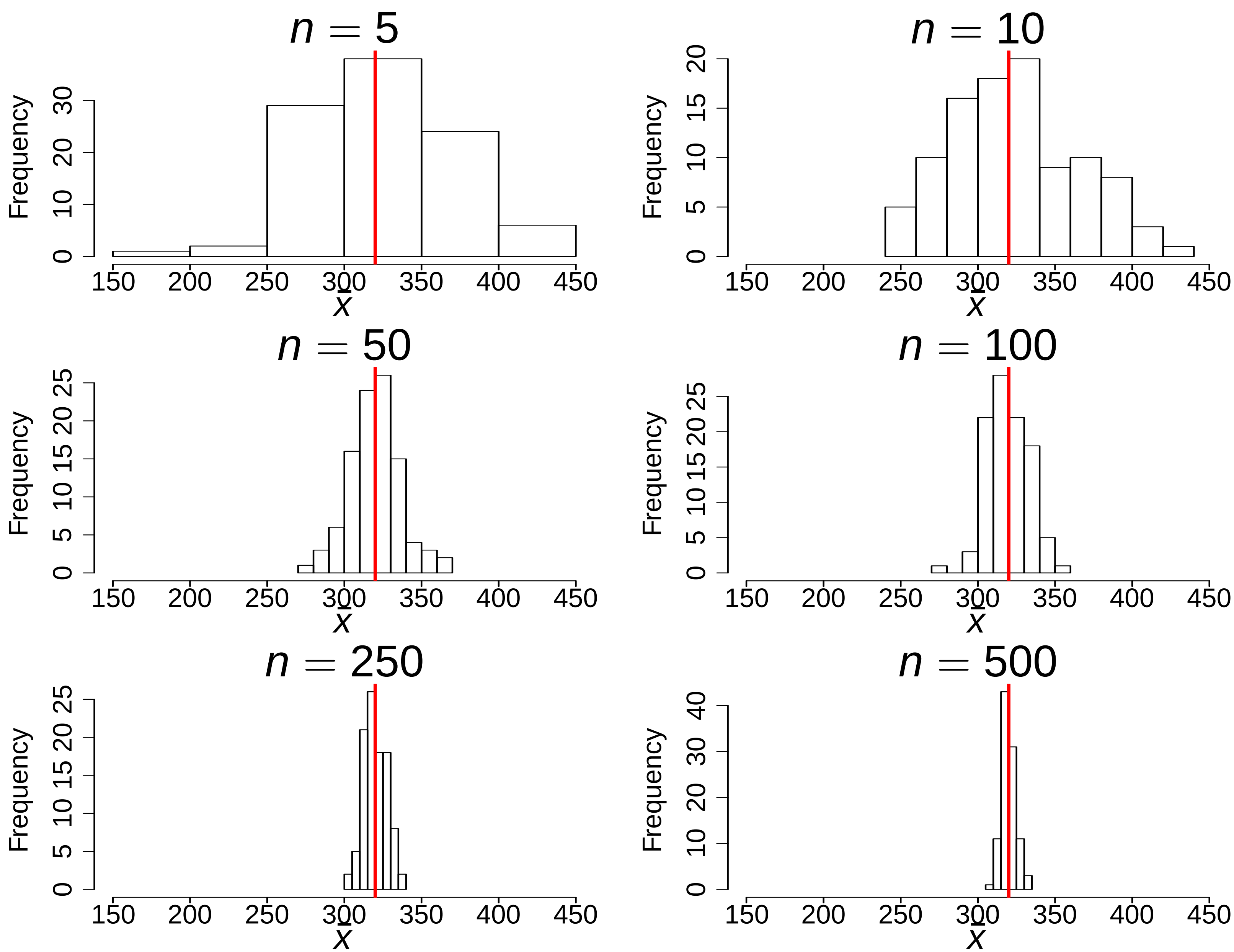
Example: *The Vintage Clothing Co.*

In fact, we could take many samples, and it's very likely that we'll get a **different value** for $\bar{x}$ each time.

It's also very *unlikely* that any of our $\bar{x}$'s will be exactly the same as the **true population mean** μ.

The graph on the next slide shows histograms of $\bar{x}$'s taken from 100 samples from our population of 1,000 stores.

Example: *The Vintage Clothing Co.*



Example: *The Vintage Clothing Co.*

You should notice two things:

1. The distribution of $\bar{X}$ looks like a **Normal distribution**;
2. As we increase the sample size (*n*), the distribution for $\bar{X}$ gets more and more **concentrated** – around the **true population value** $\mu = 320$!

In fact, what we can see in action in this graph is known as the **Central Limit Theorem**.

This is a very powerful result in Statistics which tells us about the distribution of the sample mean $\bar{X}$. We now state this formally.

The Central Limit Theorem

Suppose $x_1, x_2, \ldots, x_n$ are a random sample from *any* population, with mean μ and variance σ^2.

If n is large, then

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \text{approximately;}$$

if $x_1, x_2, \ldots, x_n$ come from a Normal distribution themselves, then this result holds for *any* n.

The Central Limit Theorem

In other words, if we were to take many samples of size n and:

- For each sample calculate the mean $\bar{x}$;
- Put all of our $\bar{x}$'s together and make a histogram of them;

then our histogram of $\bar{x}$'s will **always** be Normally distributed around the true population mean μ.

What's more, we also know about the **variability** of $\bar{X}$:

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

Thus, the standard deviation of $\bar{X}$ is $\sigma/\sqrt{n}$, and we call this the **standard error**.

6.2 Confidence intervals

The values we calculate for sample means and variances are **point estimates**.

They are single values based on a limited sample of the whole population.

Suppose that we wish to estimate the mean μ of a population. The natural estimate for μ is the sample mean $\bar{x}$.

However, as we have seen, $\bar{x}$ is never exactly equal to μ; all we really hope is that $\bar{x}$ will be close to μ.

6.2 Confidence intervals

One way of improving our inference is to construct **interval estimates**, more commonly known as **confidence intervals**.

We simply place an interval over the point estimate for μ which allows us to say (with a certain level of confidence) within what range the population mean lies.

The calculation of these intervals depends on the size of our sample (n), the level of confidence we choose, and whether or not the population variance (σ^2) is known.

6.2.1 Case 1: Known variance σ^2

We know from the CLT that, if our random sample is drawn from a Normal distribution, or if n is large (e.g. $n \geq 30$), then

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

If we initially assume we know the population variance σ^2, we can **standardise** $\bar{X}$ as we did in Chapter 5; i.e.

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}.$$

Recall that the standard Normal distribution is $Z \sim N(0, 1)$, i.e. Z has zero mean and variance (and so standard deviation) 1.

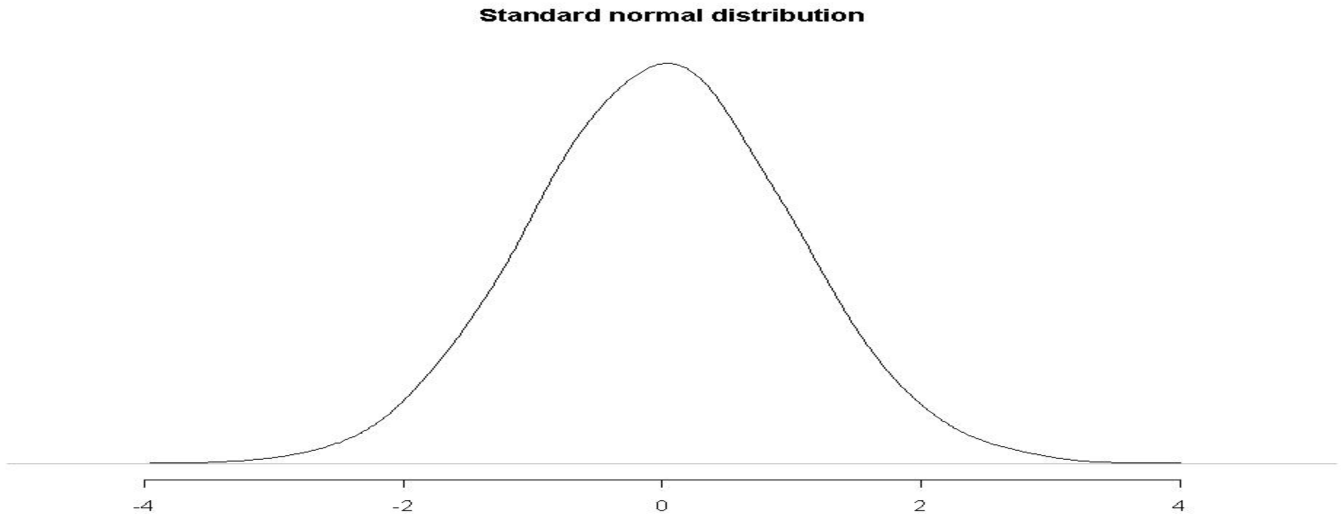
Also recall that approximately 95% of the standard normal distribution lies between -1.96 and 1.96, i.e:

6.2.1 Case 1: Known variance σ^2

- We know that (from tables)

$$\Pr(-1.96 < Z < 1.96) = 0.95;$$

We can think about this graphically:

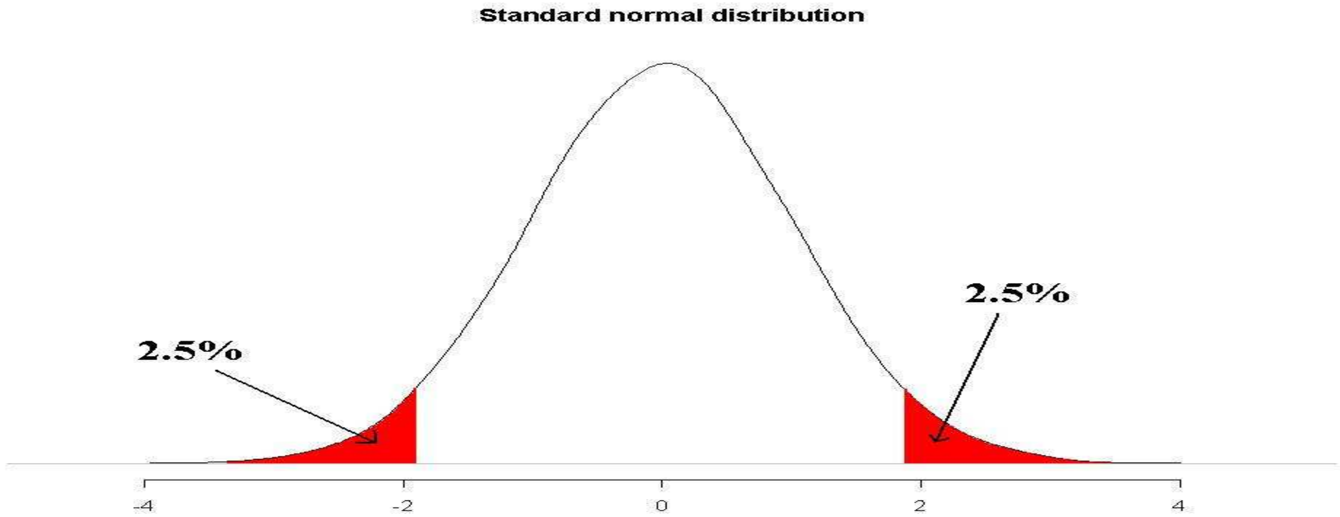


6.2.1 Case 1: Known variance σ^2

- We know that (from tables)

$$\Pr(-1.96 < Z < 1.96) = 0.95;$$

We can think about this graphically:

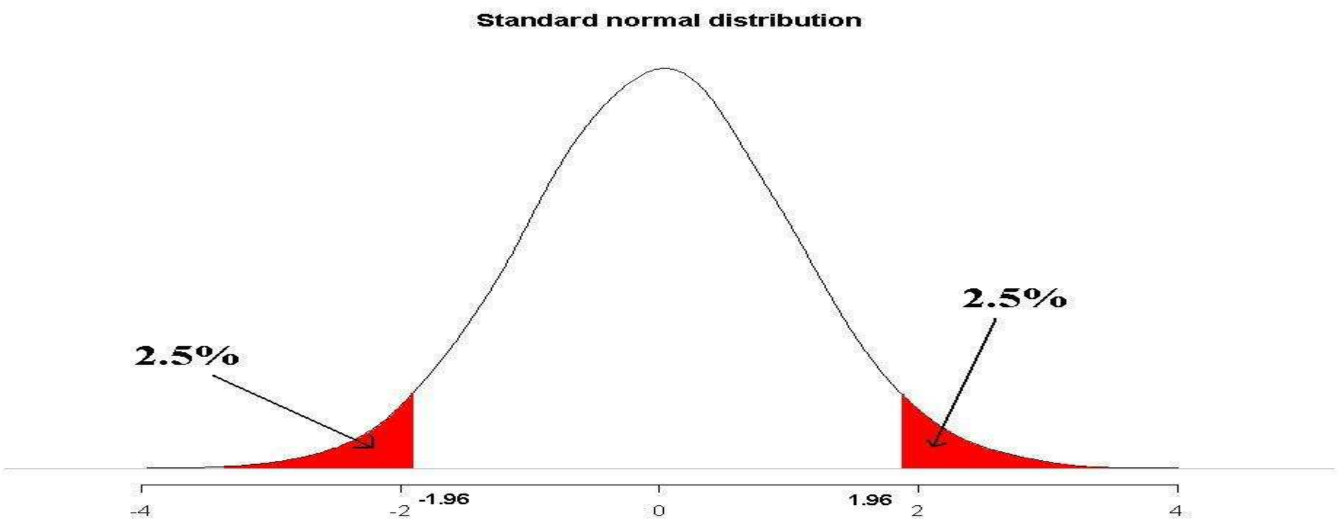


6.2.1 Case 1: Known variance σ^2

- We know that (from tables)

$$\Pr(-1.96 < Z < 1.96) \quad = \quad 0.95;$$

We can think about this graphically:

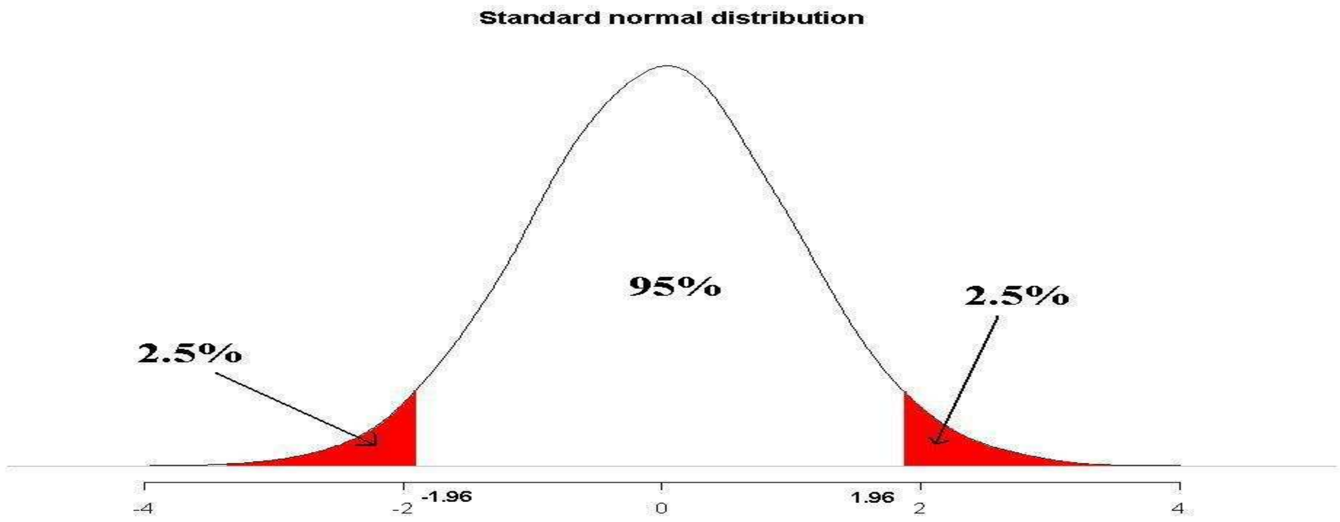


6.2.1 Case 1: Known variance σ^2

- We know that (from tables)

$$\Pr(-1.96 < Z < 1.96) \quad = \quad 0.95;$$

We can think about this graphically:



- Thus,

$$\Pr\left(-1.96 < \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} < 1.96\right) \quad = \quad 0.95;$$

6.2.1 Case 1: Known variance σ^2

Rearranging for μ gives us an expression for the **95% confidence interval for μ**:

$$\left(\bar{x} - 1.96\sqrt{\sigma^2/n} \quad , \quad \bar{x} + 1.96\sqrt{\sigma^2/n}\right);$$

thus, we can say that the two values

- $\bar{x} - 1.96\sqrt{\sigma^2/n}$ and
- $\bar{x} + 1.96\sqrt{\sigma^2/n}$

are the **lower** and **upper bounds** (respectively) of the (95%) confidence interval.

We often write this more simply as

$$\bar{x} \pm 1.96\sqrt{\sigma^2/n}.$$

6.2.1 Case 1: Known variance σ^2

Going back to *The Vintage Clothing Co.* example, this means that if we were to take 100 samples and for each one calculate a 95% confidence interval, then about 95 of these confidence intervals would “capture” the true population value $\mu = 320$.

If we wanted to be “more confident” of capturing μ, then the interval needs to be wider.

We replace the value **1.96** with

- **2.58** for a 99% interval
- **1.65** for a 90% interval

If we increase our sample size, we become more certain of our estimate $\bar{x}$ and so the width of the interval **decreases**.

Example 6.1

Geordie Sparkz are an electrical company based in Newcastle producing circuitboards for large plasma televisions.

One of their machines punches tiny holes in these curcuitboards that should be 0.5mm in diameter. A sample of 30 circuitboards off the production line is inspected; the average diameter of the holes produced by this machine, for this sample, is 0.54mm.

Assuming the machine is set to ensure a standard deviation of $\sigma =$ 0.12mm, calculate the 95% confidence interval for the population mean diameter of holes produced by this machine.

Do you think there is a real problem with this machine?

Example 6.1: Solution

We need to use the formula

$$\bar{x} \pm 1.96 \times \sqrt{\frac{\sigma^2}{n}}.$$

We have

$$n = 30, \quad \bar{x} = 0.54 \quad \text{and} \quad \sigma = 0.12.$$

This gives

$$0.54 \quad \pm \quad 1.96 \times \sqrt{\frac{0.12^2}{30}}$$

$$0.54 \quad \pm \quad 0.0429$$

This gives the 95% confidence interval as (0.497mm, 0.583mm).

No real cause for concern: specified value (0.5mm) falls within the interval.

6.2.2 Case 2: Unknown variance σ^2

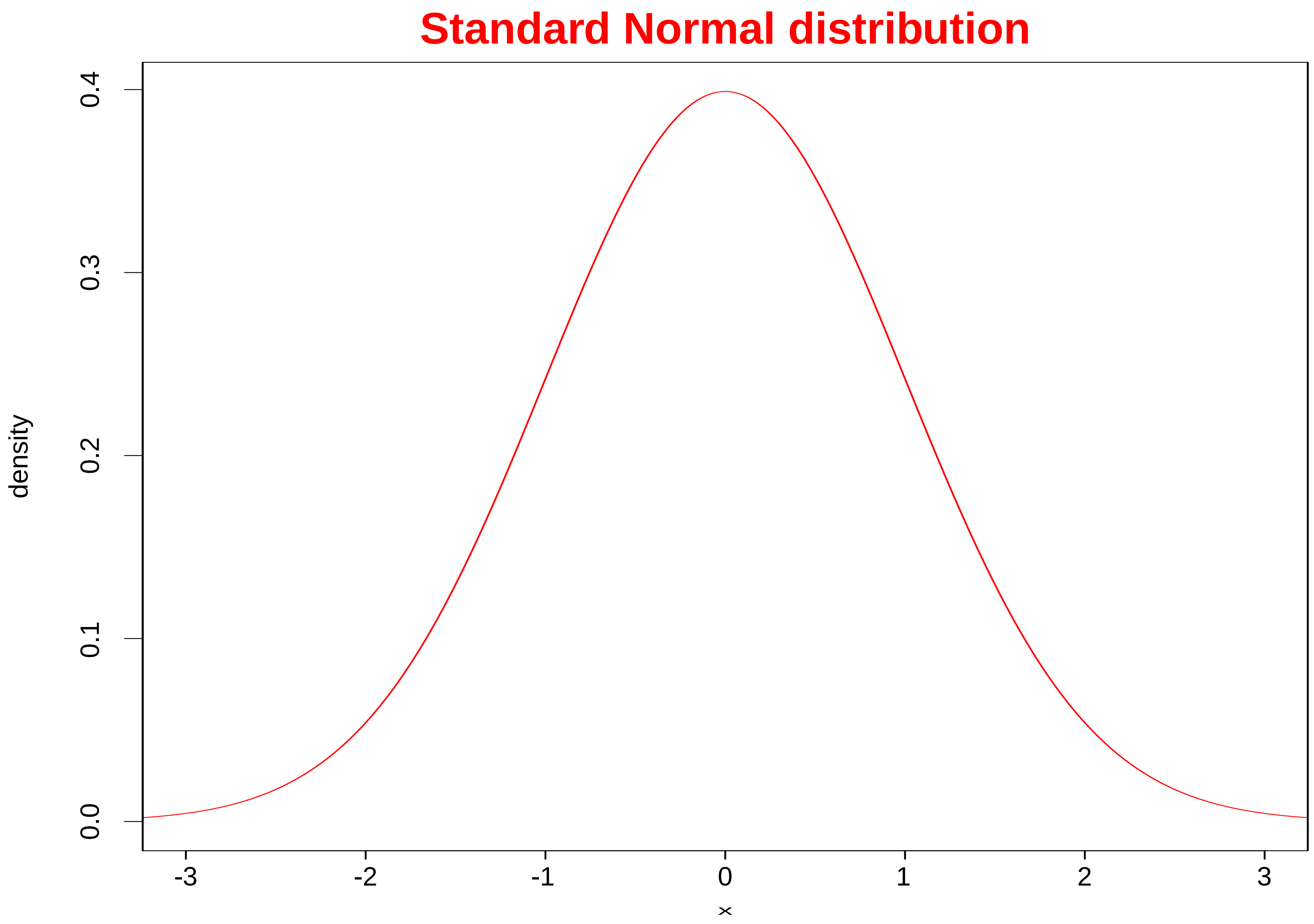
If the population variance is unknown (which is usually the case), the quantity

$$T = \frac{\bar{X} - \mu}{\sqrt{s^2/n}}$$

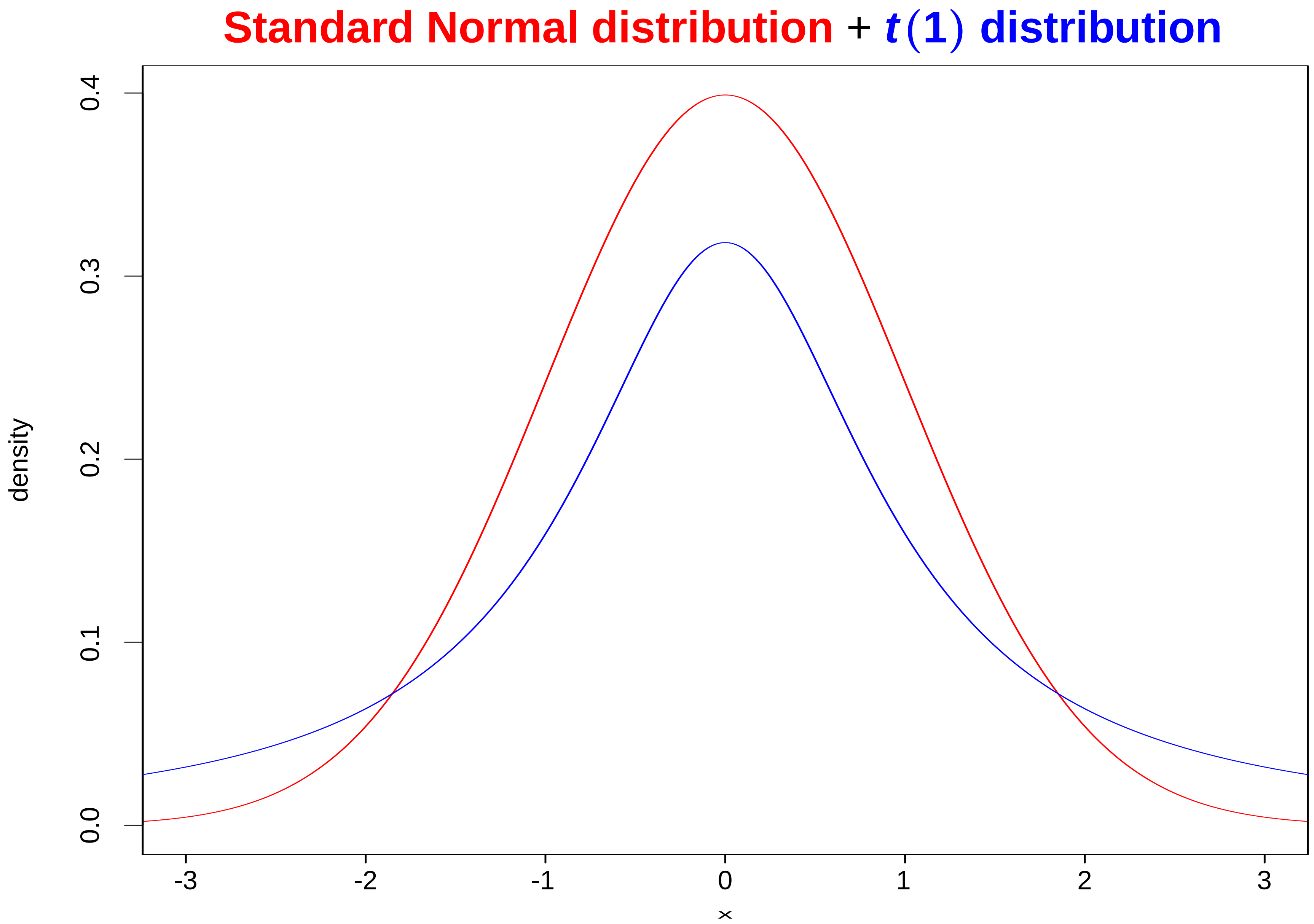
does ***not*** have a $N(0,1)$ distribution, but a **Student's *t*–distribution**.

- This is similar to the normal distribution (i.e. symmetric and bell–shaped), but is more 'heavily tailed';
- The *t*–distribution has one parameter, called the "degrees of freedom" ($\nu = n - 1$).

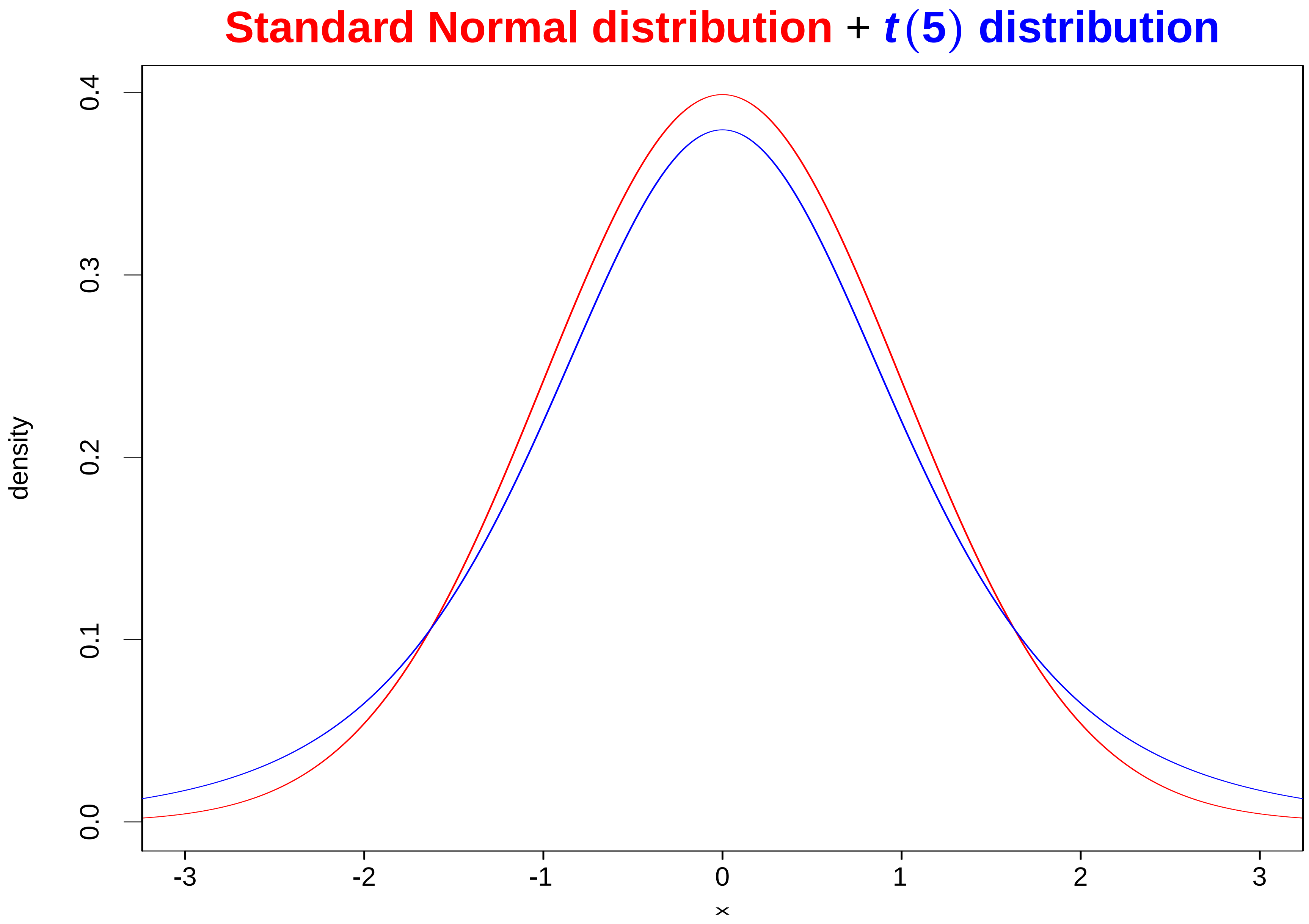
6.2.2 Case 2: Unknown variance σ^2



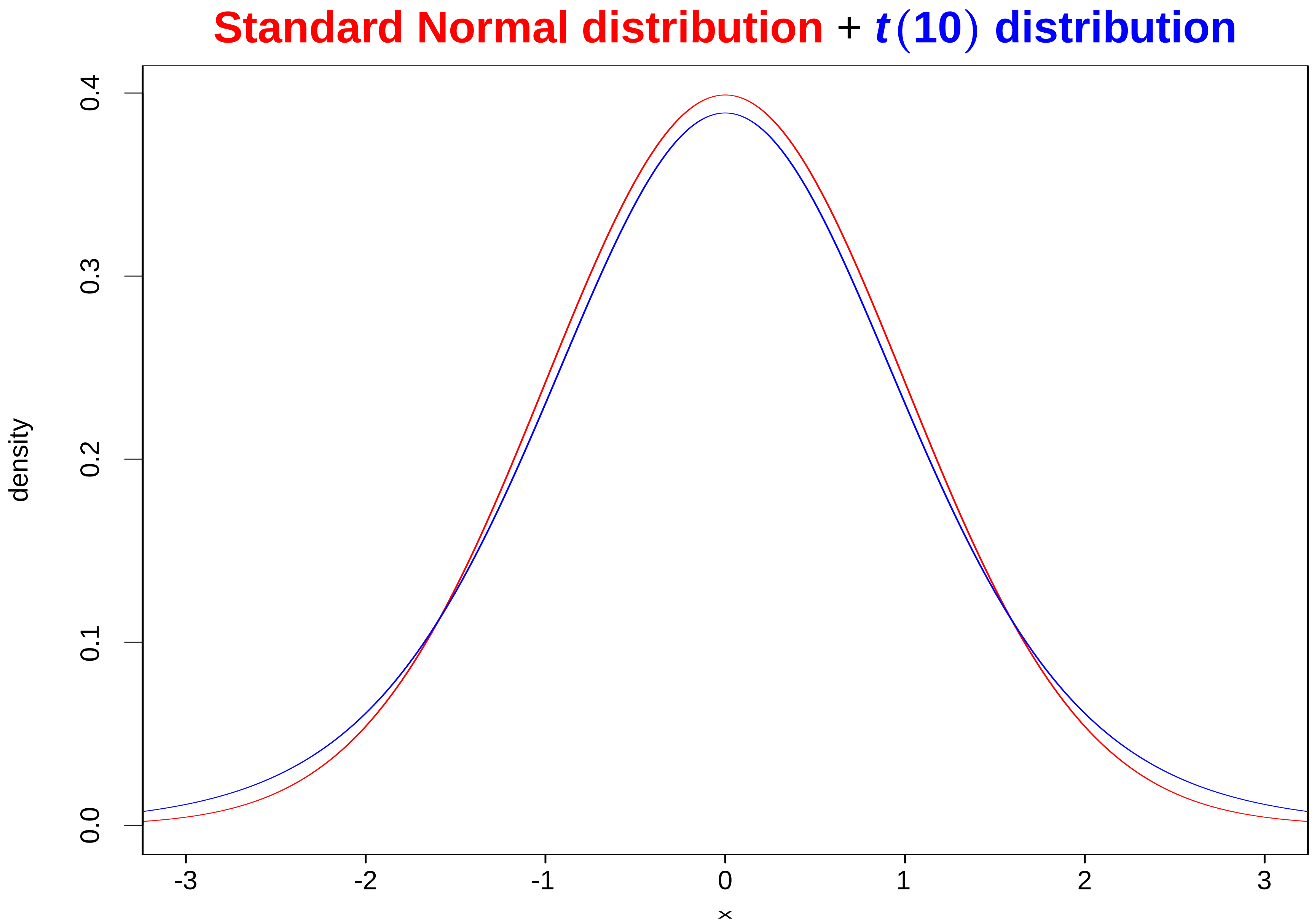
6.2.2 Case 2: Unknown variance σ^2



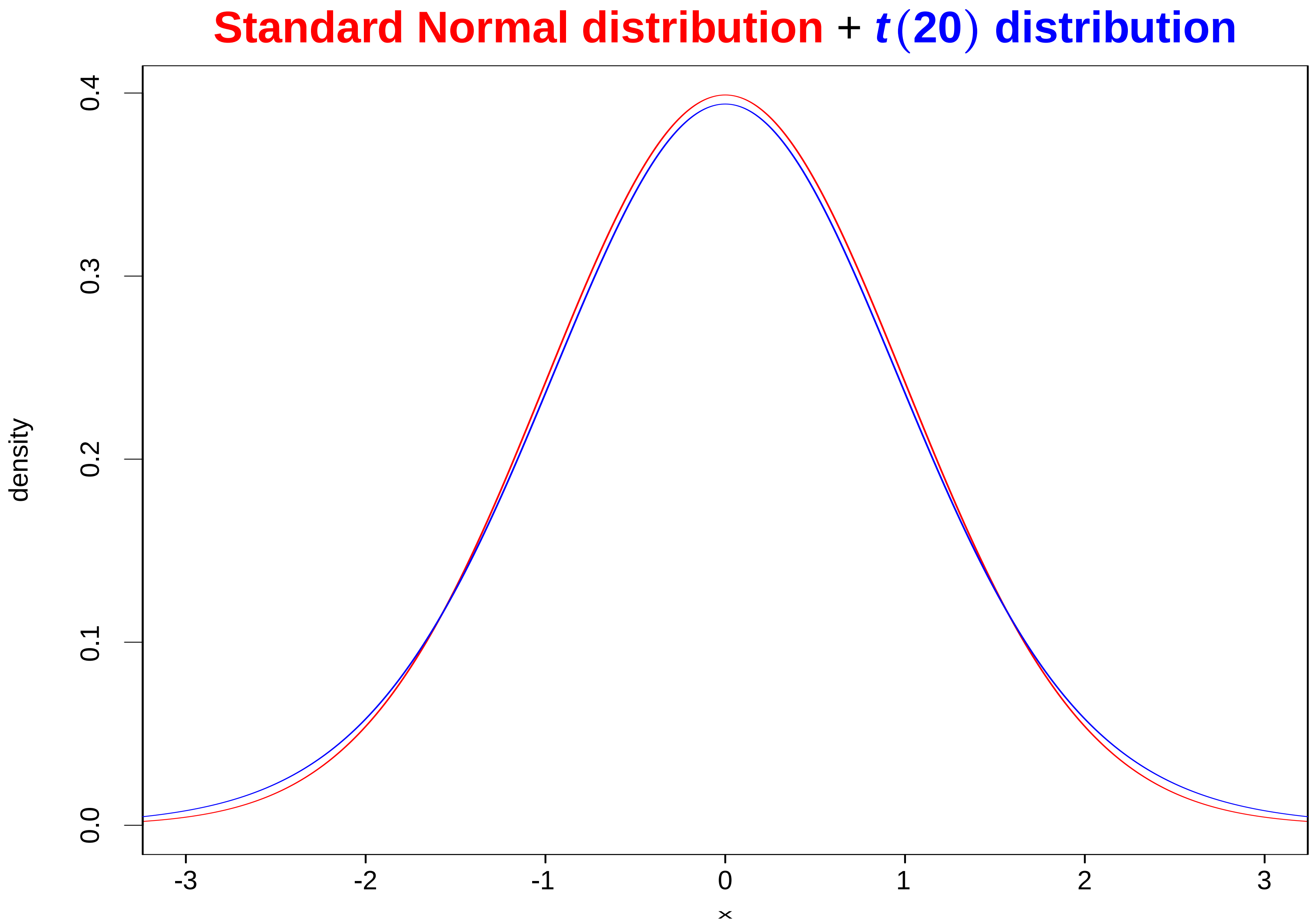
6.2.2 Case 2: Unknown variance σ^2



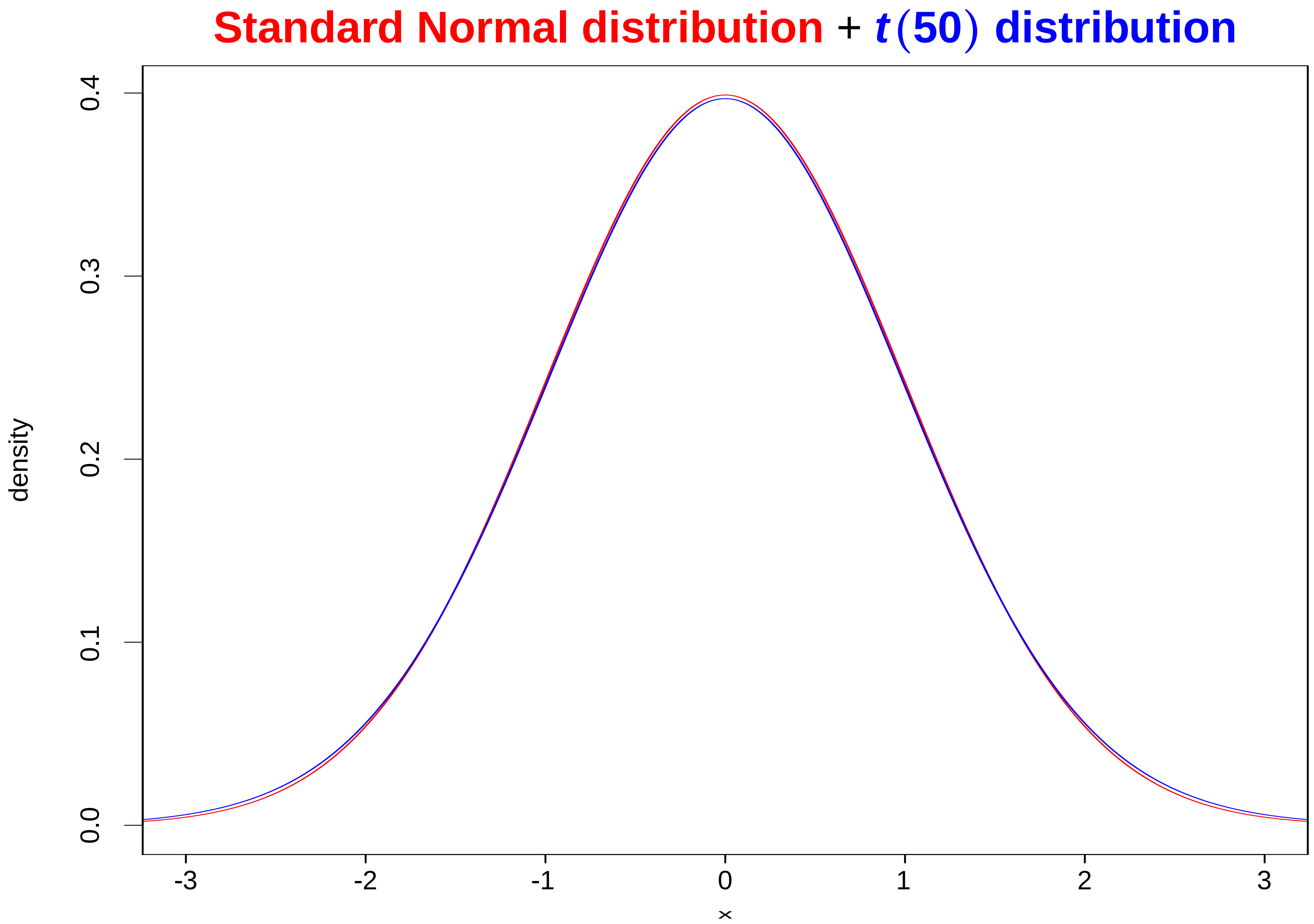
6.2.2 Case 2: Unknown variance σ^2



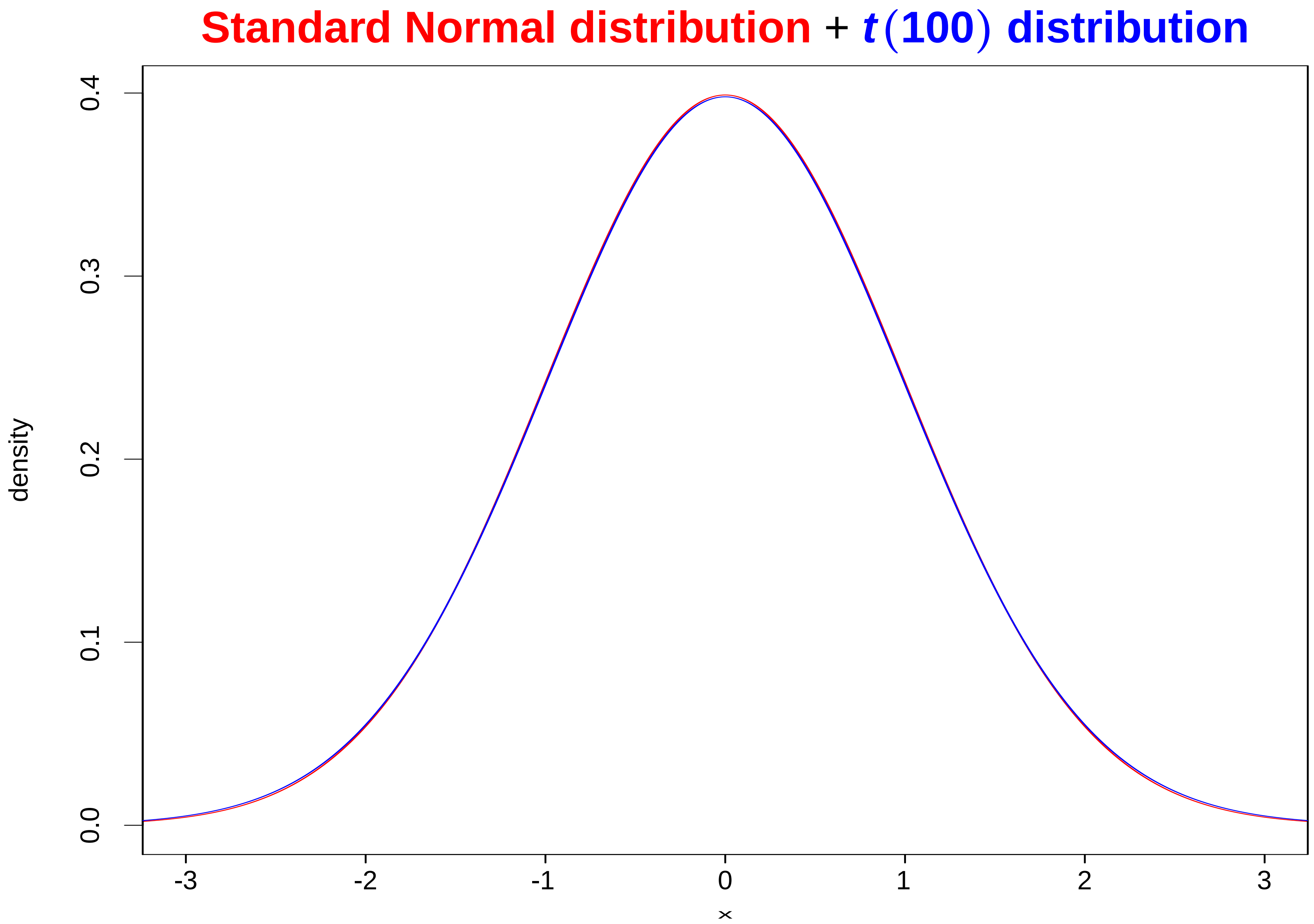
6.2.2 Case 2: Unknown variance σ^2



6.2.2 Case 2: Unknown variance σ^2



6.2.2 Case 2: Unknown variance σ^2



6.2.2 Case 2: Unknown variance σ^2

So if we don't know σ^2, the formula for the confidence interval becomes:

$$\bar{x} \pm t_p \times \sqrt{s^2/n}.$$

We find t_p from statistical tables (page 152). We read along the p column and down the ν row.

- For a 90% confidence interval, $p = 10\%$;
- For a 95% confidence interval, $p = 5\%$;
- For a 99% confidence interval, $p = 1\%$.
- The degrees of freedom, $\nu = n - 1$.

Example 6.2

A credit card company wants to determine the mean income of its card holders. It also wants to find out if there are any differences in mean income between males and females.

A random sample of 225 male card holders and 190 female card holders was drawn, and the following results obtained:

	Mean	Standard deviation
Males	£16 450	£3675
Females	£13 220	£3050

Calculate 95% confidence intervals for the mean income for males and females.

Is there any evidence to suggest that, on average, males' and females' incomes differ? If so, describe this difference.

Example 6.2: Solution

The true population variance, σ^2, is unknown (we have the **sample** variance), and so we have **case 2** and need to use the t distribution. Thus,

$$\bar{x} \pm t_p \times \sqrt{s^2/n}.$$

From the question,

$$\bar{x} = 16450, \qquad s^2 = 3675^2 = 13505625, \qquad \text{and} \qquad n = 225.$$

But what about t_p?

Example 6.2: Solution

The value t_p must be found from the table on page 152.

- Recall that the degrees of freedom, $\nu = n - 1$, and so here we have $\nu = 225 - 1 = 224$
- But the table only gives value of ν up to 29 – for higher values, we use the ∞ row
- Since we require a 95% confidence interval, we read down the $p = 5\%$ column, giving a t value of 1.96

Example 6.2: Solution

Thus, the 95% confidence interval for μ is found as

$$16450 \quad \pm \quad 1.96 \times \sqrt{13505625/225}, \qquad \text{i.e.}$$

$$16450 \quad \pm \quad 480.2.$$

So, the 95% confidence interval for **male income** is

(£15969.80, £16930.20).

Example 6.2: Solution

For females, again the true population variance, σ^2, is unknown, and so we have case 2. Thus,

$$\bar{x} \pm t_p \times \sqrt{s^2/n}.$$

Now,

$$\begin{aligned} \bar{x} &= 13220, \\ s^2 &= 3050^2 \\ &= 9302500, \qquad \text{and} \\ n &= 190. \end{aligned}$$

Example 6.2: Solution

As before, since the sample size is large we use the ∞ row of t tables to obtain the value of t_p, giving:

$$\begin{aligned} 13220 &\pm 1.96 \times \sqrt{9302500/190}, \qquad \text{i.e.} \\ 13220 &\pm 1.96 \times 221.27, \qquad \text{i.e.} \\ 13220 &\pm 433.69. \end{aligned}$$

So, the 95% confidence interval for **female income** is

(£12786.31, £13653.69).

Example 6.2: Solution

Since the 95% confidence intervals for males and females *do not overlap*, there *is* evidence to suggest that males' and females' incomes, on average, are different.

Further, it appears that male card holders earn more than women.

6.3 Hypothesis tests for one mean

We have seen that confidence intervals can be used to make inferences about population parameters.

Sometimes, you may be asked to assess whether or not a parameter takes a specific value. For example: whether the population mean $\mu = 5$.

One way of re-expressing this question is to ask whether the parameter value is plausible in light of the data.

A simple check to see whether the value is contained in a 95% confidence interval will provide an answer.

An alternative method, called a **hypothesis test**, is available. It is used extensively in reporting experimental results.

6.3 Hypothesis tests for one mean

A hypothesis test is a rule for establishing whether or not a set of data is consistent with a hypothesis about a parameter of interest.

The **null hypothesis** is a statement that a parameter has a certain value, and is usually written as H_0. For example:

- $H_0 : \mu = 2.7$
- $H_0 : \sigma^2 = 12.4$.

If the null hypothesis is not true, what alternatives are there?

Usually, the **alternative hypothesis** is written as H_1. Examples include:

- $H_1 : \mu \neq 2.7$
- $H_1 : \sigma^2 \neq 12.4$.

6.3 Hypothesis tests for one mean

Based on the information we have in our sample, we'd like to go with either the null hypothesis or the alternative hypothesis.

We might use our sample as evidence to suggest that, for example, the population mean could well be equal to 2.7; alternatively, the sample might give evidence to the contrary and suggest that the population mean is not equal to 2.7.

6.3 Hypothesis tests for one mean

Suppose you are going on holiday to Sicily in March.

A friend tells you that in March, Sicily has an average of **10 hours** sunshine a day.

On the first three days of your holiday there are **7**, **8** and **9** hours of sunshine respectively.

You consider that this is evidence that your friend is wrong.

6.3 Hypothesis tests for one mean

Thus,

- the null hypothesis would state that the average sunshine hours per day is 10
- your alternative hypothesis might state that the average sunshine hours per day *is less than* 10

Using our notation,

- $H_0 : \mu = 10$ versus
- $H_1 : \mu < 10$

6.3 Hypothesis tests for one mean

You might be tempted to go with the alternative hypothesis.

However, this sample of three days could be a fluke result – you might have chosen the most miserable period in March for years for your holiday.

Your test results are not conclusive; they only give you evidence for or against a particular belief.

6.3.1 Methodology for hypothesis testing

All hypothesis tests follow the same basic methodology, although the actual calculations may vary depending on the data available.

1. **State the null hypothesis (H_0)**
   We use a hypothesis test to throw light on whether or not this statement is true. For example, you might ask "**is the population mean equal to 10?**", or "**are the two population means equal?**"; such hypotheses are expressed in the following way:

$$
\begin{aligned}
H_0 &: \mu = 10, && \text{and} \\
H_0 &: \mu_1 = \mu_2; && \text{or maybe} \\
H_0 &: \mu = c,
\end{aligned}
$$

   Where c could be any constant.

6.3.1 Methodology for hypothesis testing

2. **State the alternative hypothesis (H_1)**
   This is the conclusion to be reached if the null hypothesis is rejected. For example, "**the population mean does not equal 10**", or even "**the population mean is less than 10**"; in our notation:

$$H_1 \quad : \quad \mu \neq 10 \qquad \text{or maybe}$$
$$H_1 \quad : \quad \mu < 10.$$

   To test for two different populations, we might say "**the two population means are different**". In our notation:

$$H_1 \quad : \quad \mu_1 \neq \mu_2.$$

6.3.1 Methodology for hypothesis testing

3. **Calculate the test statistic**
   The value calculated from the sample which is used to perform the test is called the **test statistic**.

   It usually has a similar nature to the population parameter mentioned in the null hypothesis.

6.3.1 Methodology for hypothesis testing

4. **Find the *p*–value of the test**
   The probability that such an extreme test statistic occurs, *assuming that* H_0 *is true*, is called the ***p*–value**.

   This can be found by comparing the test statistic to values from statistical tables.

5. **Reach a conclusion**

A small p–value suggests that our test statistic is unlikely to occur if H_0 is true, and so we reject H_0 in favour of the alternative H_1.

p–value	Interpretation
p is bigger than 10%	no evidence against the null hypothesis: keep H_0
p lies between 5% and 10%	*slight* evidence against H_0, but not enough to reject it
p lies between 1% and 5%	moderate evidence against H_0: reject it, and go with H_1
p is smaller than 1%	strong evidence against H_0: reject it, and go with H_1

6.3.2 Case 1: Population variance known: Example

A chain of shops believes that the average size of transactions is £130, and the population variance is known to be £900.

The takings of one branch were analysed and it was found that the mean transaction size was £123 over the 100 transactions in one day.

Based on this sample, test the null hypothesis that the true mean is equal to £130.

6.3.2 Case 1: Population variance known: Example

Since σ^2 is known (we are given that $\sigma^2 = 900$), this corresponds to case 1: population variance known (think back to confidence intervals).

We now proceed with the five steps outlined in the previous section.

Steps 1 and 2 *(hypotheses)*

Here, we state our null and alternative hypotheses. The null hypothesis is given in the question – i.e.

$$H_0 \quad : \quad \mu = £130.$$

We could test against a general alternative, i.e.

$$H_1 \quad : \quad \mu \neq £130.$$

6.3.2 Case 1: Population variance known: Example

Step 3 *(calculating the test statistic)*

When σ^2 is known, we use following test statistic

$$
\begin{aligned}
z &= \frac{|\bar{x} - \mu|}{\sqrt{\sigma^2/n}}, \qquad \text{i.e.} \\
z &= \frac{|123 - 130|}{\sqrt{900/100}} \\
&= \frac{7}{\sqrt{9}} \\
&= 2.33.
\end{aligned}
$$

6.3.2 Case 1: Population variance known: Example

Step 4 *(finding the p–value)*

Recall the **Central Limit Theorem** from Section 6.1; this tells us that the quantity

$$\frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}}$$

follows a standard Normal distribution.

Thus, the value we obtain from our test statistic formula above will be from the positive half of the standard Normal distribution.

6.3.2 Case 1: Population variance known: Example

We can therefore compare our test statistic to critical values from the standard Normal distribution to find our *p*–value, or at least a range for our *p*–value.

Remember, this is the probability of observing our data, or anything more extreme than this, if the null hypothesis is true; thus, the smaller the *p*–value, the more evidence there is **against** H_0.

6.3.2 Case 1: Population variance known: Example

Our alternative hypothesis is **two–tailed** (i.e. $\neq$ rather than $<$ or $>$), and so the critical values, from the top table on page 160, are:

Significance level	10%	5%	1%
Critical value	1.645	1.96	2.576

Our test statistic $z = 2.33$ lies between the critical values of 1.96 and 2.576, and so our p–value lies **between 1% and 5%**.

We can see this more clearly on a diagram.

6.3.2 Case 1: Population variance known: Example

Step 5 *(conclusion)*

Using table 6.2 to interpret our p–value, we see that:

- There is **moderate** evidence against H_0
- Thus, we should **reject** H_0 in favour of H_1
- It appears that the population mean transaction size **is not equal to** £130

6.3.2 Case 1: Population variance known: Example

Alternatively, since our sample mean $\bar{x} = £123$ is smaller than the proposed value of £130, we could have set up a **one–tailed** alternative hypothesis:

$$\begin{aligned} H_0 &: \mu = £130 \qquad \text{against} \\ H_1 &: \mu < £130. \end{aligned}$$

This is now a one–tailed test and the critical values from tables are:

Significance level	10%	5%	1%
Critical value	1.282	1.645	2.326

6.3.2 Case 1: Population variance known: Example

The test statistic is (as before) 2.33, which now lies “to the right” of the last critical value in the table (2.326).

Thus, our p–value is now smaller than 1%, and so, using table 6.2, we see that

- there is **strong** evidence against H_0
- We would **reject** H_0 in favour of H_1
- There is evidence to suggest that the population mean **is less than** £130

6.3.3 Case 2: Population variance unknown: Example

The batteries for a fire alarm system are required to last for 20000 hours before they need replacing.

16 batteries were tested; they were found to have an average life of 19500 hours and a standard deviation of 1200 hours.

Perform a hypothesis test to see if the batteries do, on average, last for 20000 hours.

6.3.3 Case 2: Population variance unknown: Example

Steps 1 and 2 *(hypotheses)*

Using a **one–tailed test**, our null and alternative hypotheses are:

$$
\begin{aligned}
H_0 &: \mu = 20000 \qquad \text{versus} \\
H_1 &: \mu < 20000.
\end{aligned}
$$

We use a one–tailed test because we are interested in whether the batteries are effective or not; there is no problem if they last longer than 20000 hours.

6.3.3 Case 2: Population variance unknown: Example

Step 3 *(calculating the test statistic)*

Unlike the previous example, the population variance σ^2 is unknown.

However, the **sample** standard deviation is given, based on a sample of size 16, and so we need to use a slightly different test statistic.

In fact, we do what we did last lecture when we were constructing confidence intervals – i.e. we replace σ^2 with s^2 and then use tables of values from Student's t distribution instead of the standard Normal distribution.

6.3.3 Case 2: Population variance unknown: Example

Thus, the test statistic is given by

$$
\begin{aligned}
t &= \frac{|\bar{x} - \mu|}{\sqrt{s^2/n}} \\
&= \frac{|19500 - 20000|}{\sqrt{1200^2/16}} \\
&= \frac{500}{\sqrt{1440000/16}} \\
&= 1.667.
\end{aligned}
$$

6.3.3 Case 2: Population variance unknown: Example

Step 4 *(finding the p–value)*

Since σ^2 is unknown, we use *t*–distribution tables to obtain a range for our *p*–value.

The degrees of freedom, $\nu = n - 1 = 16 - 1 = 15$, and under a one–tailed test this gives the following critical values:

Significance level	10%	5%	1%
Critical value	1.341	1.753	2.602

Our test statistic of $t = 1.667$ lies between the critical values of 1.341 and 1.753, and so the corresponding *p*–value lies between 5% and 10%.

6.3.3 Case 2: Population variance unknown: Example

Step 5 *(conclusion)*

Using table 6.2 to interpret our p–value, we see that there is

- only **slight** evidence against the null hypothesis
- this is not enough grounds to reject it, so we **retain** H_0
- There is insufficient evidence to suggest there is a problem with these batteries.

6.4 Testing two means

If we have **two** independent random samples from **two** populations, we can compare the two sample means (*c.f.* comparing one sample mean to a *proposed value* in the one–sample case).

We use the same framework for hypothesis testing as for the one–sample test; however, the calculations required for the test statistic are slightly different. There are two situations to consider here:

- Are **both** population variances (σ_1^2 **and** σ_2^2) **known**?
- Are **both** population variances **unknown**?

6.4.1 Both population variances (σ_1^2 and σ_2^2) known

1. **State the null hypothesis**
   This time, the null hypothesis is
   $$H_0 \quad : \quad \mu_1 = \mu_2,$$
   i.e. the two population means are equal.

2. **State the alternative hypothesis**
   We usually test against the (**two–tailed**) alternative:
   $$H_1 \quad : \quad \mu_1 \neq \mu_2,$$
   i.e. the population means *are not* equal. However, we might use the **one–tailed** alternatives:
   $$H_1 \quad : \quad \mu_1 > \mu_2, \qquad \text{or}$$
   $$H_1 \quad : \quad \mu_1 < \mu_2.$$

6.4.1 Both population variances (σ_1^2 and σ_2^2) known

3. **Calculate the test statistic**
   The test statistic for a two–sample test (when both population variances are known) is

$$Z = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Both population variances (σ_1^2 and σ_2^2) known

4. **Find the *p*–value**
   This is found from statistical tables; since, in this case, both population variances σ_1^2 and σ_2^2 are *known*, we refer to standard normal tables. As before, we find a range for our *p*–value by comparing our test statistic to the 10%, 5% and 1% critical values.

5. **Form a conclusion**
   Exactly the same again! Use table 6.2 to help you decide what to do! Word your conclusions in the context of the original question.

Example 6.3

Before a training session for call centre employees a sample of 50 calls to the call centre had an average duration of 5 minutes, whereas after the training session a sample of 45 calls had an average duration of 4.5 minutes.

The population variance is **known** to have been 1.5 minutes before the course and 2 minutes afterwards.

Has the course been effective?

Example 6.3: Solution

Steps 1 & 2 *(hypotheses)*

$$H_0 \quad : \quad \mu_B = \mu_A \qquad \text{versus}$$

$$H_1 \quad : \quad \mu_B \neq \mu_A$$

Example 6.3: Solution

Step 3 *(Test statistic)*

We use

$$
\begin{aligned}
Z &= \frac{|\bar{X}_B - \bar{X}_A|}{\sqrt{\frac{\sigma_B^2}{n_B} + \frac{\sigma_A^2}{n_A}}} \\
&= \frac{|5 - 4.5|}{\sqrt{\frac{1.5}{50} + \frac{2}{45}}} \\
&= \frac{0.5}{\sqrt{0.03 + 0.0444}} \\
&= 1.833
\end{aligned}
$$

Example 6.3: Solution

Step 4 *(p–value)*

Since this is case 1 (both population variances known), we use Normal tables (top of page 160).

Recall that we have a **two–tailed** alternative, and so we get

Significance level	10%	5%	1%
Critical value	1.645	1.96	2.576

Our test statistic $z = 1.833$ shows that our p–value is between 5% and 10%.

Example 6.3: Solution

Step 5 *(Conclusion)*

From Table 6.2,

- We have **slight** evidence against H_0
- This is not enough to reject and so we **retain** H_0
- There is insufficient evidence to suggest the training course has been effective!

6.4.2 σ_1^2 and σ_2^2 unknown

In the more likely situation where the population variances are unknown, the test statistic becomes

$$T = \frac{|\bar{X}_1 - \bar{X}_2|}{S \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

where S is a “pooled standard deviation”, and is found as

$$S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}.$$

Like before, we have to use t–tables to obtain our critical value; the degrees of freedom is now found as $\nu = n_1 + n_2 - 2$.

Example 6.4

A company is interested in knowing if two branches have the same level of average transactions.

The company sample a small number of transactions and calculates the following statistics:

$$\begin{array}{l|lll} \text{Shop 1} & \bar{x}_1 = 130 & s_1^2 = 700 & n_1 = 12 \\ \text{Shop 2} & \bar{x}_2 = 120 & s_2^2 = 800 & n_2 = 15 \end{array}$$

Test whether or not the two branches have (on average) the same level of transactions.

Example 6.4: Solution

Steps 1 and 2 *(hypotheses)*

Our null and alternative hypotheses are:

$$H_0 \quad : \quad \mu_1 = \mu_2 \qquad \text{versus}$$
$$H_1 \quad : \quad \mu_1 \neq \mu_2.$$

Step 3 *(Test statistic)*

Since both population variances are unknown (only the *sample* values are given), the test statistic is

$$T = \frac{|\bar{X}_1 - \bar{X}_2|}{S \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

Example 6.4: Solution

First the **pooled standard deviation**:

$$
\begin{aligned}
S &= \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} \\
&= \sqrt{\frac{(12-1)\times 700+(15-1)\times 800}{12+15-2}} \\
&= \sqrt{\frac{11\times 700+14\times 800}{25}} \\
&= \sqrt{\frac{7700+11200}{25}} \\
&= \sqrt{\frac{18900}{25}} \\
&= \sqrt{756}=27.495.
\end{aligned}
$$

Example 6.4: Solution

Now we put this into our formula for t:

$$
\begin{aligned}
t &= \frac{|130 - 120|}{27.495 \times \sqrt{\frac{1}{12} + \frac{1}{15}}} \\
&= \frac{10}{27.495 \times \sqrt{0.15}} \\
&= \frac{10}{10.649} \\
&= 0.939.
\end{aligned}
$$

Example 6.4: Solution

Step 4 *(p–value)*

Since both population variances are unknown, we use t–tables to obtain our critical values.

The degrees of freedom, $\nu = n_1 + n_2 - 2$, i.e. $\nu = 12 + 15 - 2 = 25$.

Under a two–tailed test, and using t–tables (bottom table, page 160), we get the following critical values:

Significance level	10%	5%	1%
Critical value	1.708	2.060	2.787

Our test statistic $t = 0.939$ lies to the left of the first critical value, and so our p–value is bigger than 10%.

Example 6.4: Solution

Step 5 *(Conclusion)*

Using Table 6.2,

- We have **no** evidence against H_0
- Therefore, we **retain** H_0
- There is no significant difference between the average level of transactions at the two shops.