Aims of this lecture

Inference for proportion and mean

1. What is Hypothesis Testing?

- Hypothesis Testing on Proportion

2. Hypothesis Testing on Mean

Aim 1 Statistical Inference: in general

Select from

Population | Sample

Represent for

property of

Parameter

Statistical Inference:
- Hypothesis testing
- Estimation:
  1. Point estimation
  2. Confidence interval

property of

Statistic

examples:		examples:
μ	mean	$\bar{x}$
σ	standard deviation	s
p	proportion	$\hat{p}$

6 Steps in carrying out a hypothesis test: in general

1. State the hypotheses
2. Calculate the test statistic
3. Determine the sampling distribution of the test statistic in (2)
4. Find the p-value based on (3)
5. Make a decision based on (4)
6. State your conclusion in the context of your specific setting.

What is a Hypothesis?

- A statement about the population parameter in question.
- We generally use two contradictory statements (hypotheses)
  - the null hypothesis, H_0
  - the alternative hypothesis, H_A (or H_1)

Hypotheses: two contradictory statements

- Null hypothesis (H_0): Usually represent the status quo, the no effect is present or the prior belief.

- Alternative hypothesis (H_A): Simply states the case that the parameter differs from its null (H_0) in a specific direction (one sided) or in either direction (two-sided).

- Usually both H_0 and H_A are stated in terms of the population parameters .

Example 1: Optus customer

- A sample of university student mobile phone users were asked about their service provider. 18 of the 33 students surveyed were Optus customers.
- Optus claims that the proportion of Optus customers among university student mobile phone users is 50%.

(i) Estimate the (population) proportion of university student mobile phone users who are Optus customers.
(ii) Does the sample provide evidence to support Optus' claim?

Understanding the questions

- (i) Point estimate: sample proportion for population proportion
- (ii) Hypothesis Testing for a population proportion – two sided test

AIM 1 Step 1: State the null and alternate hypotheses

Suppose that, p, be the population proportion of Optus customers among university student mobile phone users

A point estimate of p is the sample proportion $\hat{p}$,
Two-sided hypothesis testing:

STEP 1 H_0: $p=0.5$ H_A: $p\neq0.5$

In words:
H_o: the (population) proportion of university student mobile phone users who are Optus customers is 50%.

H_A: the (population) proportion of university student mobile phone users who are Optus customers is not 50% (ie >50% or <50%)

Step 2: Test Statistic

- In hypothesis tests the test statistic summarises the differences between the observed and expected data.

- It is a random variable with a distribution that we know.

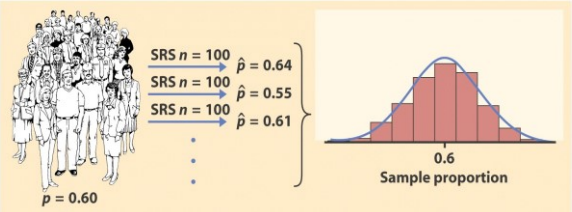
What is a sampling distribution? (For Step 3)

- The **sampling distribution of a statistic** is the distribution of all possible values taken by **the test statistic** when all possible samples of a fixed size *n* are taken from the population.
- It is a theoretical idea—we do not actually build it.
- The sampling distribution of a statistic is the **probability distribution** of that statistic.

Sampling variability

- Each time we take a random sample from a population, we are likely to get a different set of individuals and calculate a different statistic. This is called **sampling variability.**

- If we take a lot of random samples of the same size from a given population, the variation from sample to sample—the **sampling distribution**—will follow a predictable pattern.



Binomial distribution in statistical sampling

- A population contains a proportion *p* of successes. If the population is much larger than the sample, the count *X* of successes in an Simple Random Sample (SRS) of size n has approximately the binomial distribution *B*(*n*, *p*).
- The *n* observations will be nearly independent when the size of the population is much larger than the size of the sample.
- As a rule of thumb, the **binomial sampling distribution for counts** can be used when the population is at least 20 times as large as the sample.

Sample proportions

- The proportion of "successes" can be more informative than the count. In statistical sampling the sample proportion of successes, $\hat{p}$, is used to estimate the proportion *p* of successes in a population.

- For any SRS of size *n*, the sample proportion of successes is: $\hat{p} = \dfrac{\text{count of successes in the sample}}{n} = \dfrac{X}{n}$

▫ **Example 2.** In an SRS of 50 students in an undergrad class, 10 are Hispanic:

$\hat{p}$ = (10)/(50) = 0.2 *(proportion of Hispanics in sample)*

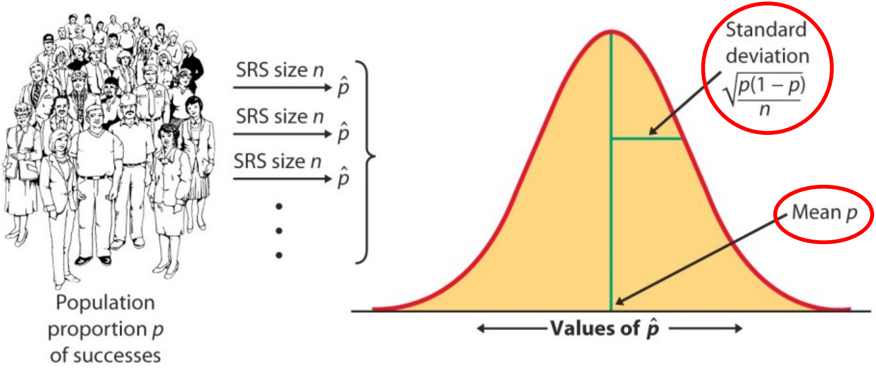
▫ **Example 3.** The 30 subjects in an SRS are asked to taste an unmarked brand of coffee and rate it "would buy" or "would not buy." Eighteen subjects rated the coffee "would buy."

$\hat{p}$ = (18)/(30) = 0.6 *(proportion of "would buy")*

Sampling distribution of the sample proportion

The sampling distribution of $\hat{p}$ is never exactly normal. But as the sample size increases, the sampling distribution of $\hat{p}$ becomes approximately normal.

The normal approximation is most accurate for any fixed *n* when *p* is close to 0.5, and least accurate when *p* is near 0 or near 1.



Normal approximation to Binomial

If n is large, and p is not too close to 0 or 1, the binomial distribution can be approximated by the Normal distribution $N(\mu = np, \sigma^2 = np(1-p))$.

Practically, the Normal approximation can be used when both $np \geq 10$ and $n(1-p) \geq 10$.

If X is the count of successes in the sample and $\hat{p} = X/n$, the sample proportion of successes, their sampling distributions for large n, are:

X approximately $N(\mu = np, \sigma^2 = np(1-p))$

$\hat{p}$ **is approximately $N(\mu = p, \sigma^2 = p(1-p)/n)$**

Steps 2 and 3: Calculate the test statistic and the sampling distribution

Test statistic – is z–test and its value is calculated as

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$$

Solutions Example 1: Optus Customer Steps 2 and 3

For the data, we have n = 33 trials, X = 18 successes.
A point estimate of the population proportion of Optus customers among university student mobile phone users, *p*, is the sample proportion,

$$\hat{p} = X/n = 18/33 = 0.5455$$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$$

We have p0=0.5 (under H_0),
Two-sided hypothesis testin

STEP 1 H_0: $p = p_0 = 0.5$ H_A: $p \neq 0.5$ (two-sided)

STEP 2 The test statistic is

$$z = \frac{0.5455 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{33}}} = \frac{0.0455}{0.087} = 0.5228$$

STEP 3 The sampling distribution: **z ~ N(0,1)**

Step 4: p-value

The test statistic helps us find the probability of getting an outcome:

"*as extreme as, or more extreme than, the actual observed outcome*".

"extreme" means "far" from what we would expect if the null was true.

What is a p-value?

DEFINITION: The probability ***of getting the observed value of the test statistic or a more extreme value if the null is true.***

The p-value can be interpreted as the strength of the evidence provided by the observed data against H_0 .

p-value must always be between 0 and 1.
small p-value: strong evidence against H_0
large p-value: weak evidence against H_0

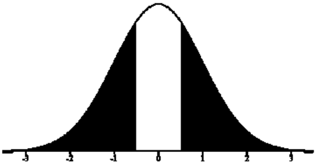
Step 4: Calculate the *p-value* for Optus data

Test statistic: $z = 0.5228$

STEP 4 As this is a two-sided test [from Ha]:

$$\begin{aligned}\text{p-value} &= P(|Z| > 0.5228) \\ &= P(Z < -0.5228) + P(Z > 0.5228) \\ &= 2P(Z < -0.5228) = 2(0.3006) \\ &= 0.6012 \text{ (The table or Excel)}\end{aligned}$$

Excel:
=NORM.S.DIST(-0.5228,TRUE)
=0.3006



So far we've discussed:

STEP 1 Hypotheses – H_o and H_A
STEP 2 Test statistic
STEP 3 Sampling distribution of the test statistic
STEP 4 p-value

Next we need to decide whether or not to reject H_o.

Step 5: Decision

- If the p-value is small, reject H_o
- If the p-value is large, do not reject H_o

Decision based on a Significance Level

- We can also compare the obtained *p*–value with a fixed value that the researcher regards as decisive.
- Meaning that you need to state in advance how much evidence against the null you will require to reject the null.
- This fixed value is called the significance level and usually known as α

Decision based on a significance level

The decision requires us to compare the p-value we found from the test statistic with the significance level α .

If p-value > significance level α

do not reject H_o.

If p-value $\leq$ significance level α

reject H_o.

If the p-value is as small or smaller than α , we say that the data is statistically significant at level α

Step 5 – Decision for Optus example

As the p-value is large (.6012) we **do not reject *H_o*.**

Based on a significance level of 5%:

As the p- value (.6012) is greater than □ (.05),

we do not reject *H_o*.

Steps 5 and 6: Decision and Conclusion based on large p-value

As the p–value (.6012) is large, we do not reject H_o and conclude that the sample is compatible with Optus' claim.

(Conclusions should be stated in the context of the problem posed)

Steps 5 and 6: Decision and Conclusion based on a level of significance

At 5% significance level we do not reject H_o and conclude that the sample is compatible with Optus' claim.

(Conclusions should be stated in the context of the problem posed)

Summary:

Large-sample significance test for a Population Proportion: Two-sided test

STEP 1 H_0: $p=p_0$ H_A: $p \neq p_0$

STEP 2 The test statistic is: $$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$$

STEP 3 Normal approximation to Binomial:

$$Z \sim N(0,1)$$

STEP 4 p-value = $P(|Z| > z) = P(Z < -z) + P(Z > z)$

STEP 5 Decision

STEP 6 Conclusion

Solutions Example 1 Optus:

All 6 steps

For the data, we have n = 33 trials, X = 18 successes.

A point estimate of the population proportion of Optus customers among university student mobile phone users, p, is the sample proportion,

$\hat{p}$ = X/n =18/33= 0.5455

Two-sided hypothesis testing:

STEP 1 H_0: p=0.5 H_A: $p \neq 0.5$

STEP 2 The test statistic is $$z = \frac{0.5455 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{33}}} = 0.5228$$

STEP 3 **z ~ N(0,1)**

STEP 4 p-value = P(|Z |> 0.5228) = P(Z < -0.5228) + P(Z > 0.5228) 0.6012 (The table or R)

STEPS 5 and 6 Based on this large p-value of 60%, we would not reject H_0 and we conclude that the sample is compatible with Optus' claim.

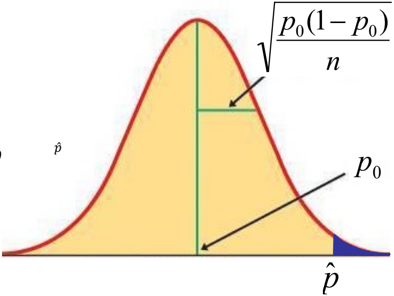
Significance test for p

The sampling distribution for $\hat{p}$ is approximately normal for large sample sizes and its shape depends solely on p and n.

Thus, we can easily test the null hypothesis:
H_0: $p = p_0$ (a given value we are testing).

If H_0 is true, the sampling distribution is known
→
The likelihood of our sample proportion given the null hypothesis depends on how far from p_0 our $\hat{p}$ is in units of standard deviation.

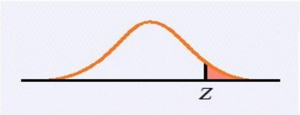
$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$$



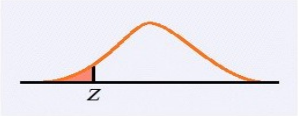
This is valid when both expected counts—expected successes np_0 and expected failures $n(1 - p_0)$—are each 10 or larger.

P-values and one or two sided hypotheses

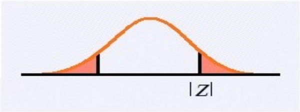
$H_a: p > p_0$ is $P(Z \geq z)$



$H_a: p < p_0$ is $P(Z \leq z)$



$H_a: p \neq p_0$ is $2P(Z \geq |z|)$



And as always, if the p-value is as small or smaller than the significance level α, then the difference is statistically significant and we reject H_0.

Example 4: Topic Notes

Suppose that Optus has a new claim that the proportion of Optus customers among university student mobile phone users is greater than 50%.

Solutions

One-sided hypothesis testing:

STEP 1 H_0: p=0.5 ***H_A: p > 0.5***

STEP 2 The test statistic is $z = \frac{0.5455 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{33}}} = 0.528$

STEP 3 Z ~ N(0.1)

STEP 4 p-value = P(Z > 0.528) = 1 - P(Z < 0.5228)

= 1 – 0.6994 = 0.3006 (Normal Table or Excel)

STEPS 5 and 6 As the p-value of 30% is large, we would not reject H_0 and therefore conclude that based on this sample, one cannot claim that the proportion of Optus customers among university student mobile phone users is greater than 50%.

Large-sample significance test for a Population Proportion: One-sided test

STEP 1 H_0: $p = p_0$ H_A: $p > p_0$ **(look at H_A)**

STEPS 2 and 3 The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$$

STEP 4 p–value

p–value = P(Z > z) **(look at H_A)**

STEP 5 Decision

STEP 6 Conclusion

Large-sample significance test for a Population Proportion: One-sided test

STEP 1 H_0: $p = p_0$ H_A: $p < p_0$ **(look at H_A)**

STEPS 2 and 3 The test statistic and sampling distribution

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$$

STEP 4 p-value

p-value = P(Z < z) **(look at H_A)**

STEP 5 Decision

STEP 6 Conclusion

Testing a population mean: one sample *t* test

When σ is known – Z test

When σ is unknown – *t*–test

σ is the population standard deviation

When σ is known

Notations:

- μ is the population mean
- σ is the population standard deviation
- n is the sample size
- $\overline{x}$ is the sample mean

Z-test:

Test statistic:

$$z = \frac{(\bar{x} - \mu)}{\sigma / \sqrt{n}}$$

which follows the standard Normal distribution, N(0,1)

When σ is unknown

Notations:

- μ is the population mean
- σ is the population standard deviation
- *s* is the sample standard deviation
- *n* is the sample size
- $\bar{x}$ is the sample mean

t–test:

Test statistic:

$$t = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

which follows Student's *t*–distribution with df = n–1

Student’s *t*-Distribution

If **s** is unknown (which is usually the case) we approximate it with ***s*** (standard deviation of sample) and use

Student’s t-distribution

instead of the Normal distribution

Properties of the *t*-distribution

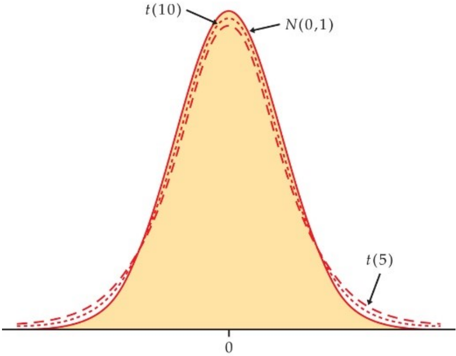
Symmetric

More widely dispersed than N(0,1)

Shape depends on sample size *n*

(n-1) degrees of freedom in this case.

For *n* very large, *t* is close to N(0,1)



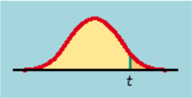
Hypothesis testing: one sample t-test

The ***p-value*** is the probability, if H_0 is true, of randomly drawing a sample like the one obtained or more extreme, in the direction of H_A.

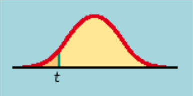
The p-value is calculated as the corresponding area under the curve, one-tailed or two-tailed depending on H_A:

One-sided (one-tailed)

$H_a\!: \mu > \mu_0 \Rightarrow P(T \geq t)$



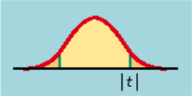
$H_a\!: \mu < \mu_0 \Rightarrow P(T \leq t)$



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

Two-sided (two-tailed)

$H_a\!: \mu \neq \mu_0 \Rightarrow 2P(T \geq |t|)$



6 Steps of hypothesis testing: t-test one sample mean (two-sided)

1. State the hypotheses:

$H_0: \mu = \mu_0 \quad H_A: \mu \neq \mu_0$

2. Calculate the test statistic

$$t = \frac{\bar{x} - \mu}{\left(s/\sqrt{n}\right)}$$

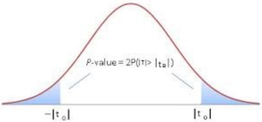
Notations:

- μ is the population mean
- μ_0 is the value of the population mean
- s is the sample standard deviation
- n is the sample size
- $\bar{x}$ is the sample mean

3. Determine the sampling distribution of the test statistic in (2)

$$t = \frac{\bar{x} - \mu}{\left(s/\sqrt{n}\right)} \sim T_{(n-1)}$$

4. Find the p-value based on (3): $P(|T(df=n-1)| > t)$
5. Make a decision based on (4)
6. State your conclusion in the context of your specific setting.



Example 5: Seedlings

A forest ecologist studying regeneration of rainforest communities in gaps caused by large trees falling during storms, read that stinging tree seedlings will grow 1.5m/year in direct sunlight in such gaps.
In the gaps in her study plot she identified nine specimens of this species and measured them in 1998 and again 1 year later.
Does her data support the published contention that seedlings for this species will average 1.5m of growth per year in direct sunlight?

Step 1 : Hypotheses (two-tailed test)

H_0: $\mu = 1.5$

H_A: $\mu \neq 1.5$

Steps 2 and 3: Test statistic

How far away is the sample mean from the required value (H_0)?

What is the probability of being this far away if the true mean **really is** the required value (1.5 in this case)?

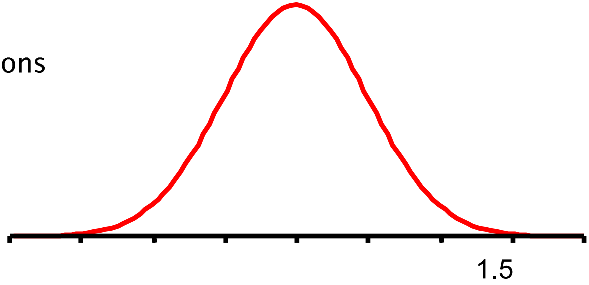
CLT tells us that for large samples,

$$\bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

approximately.

Therefore we could calculate a probability to estimate how likely this sample mean is.

We could also calculate how many standard deviations away the sample mean is (from 1.5)



Steps 2 and 3: The sampling distribution

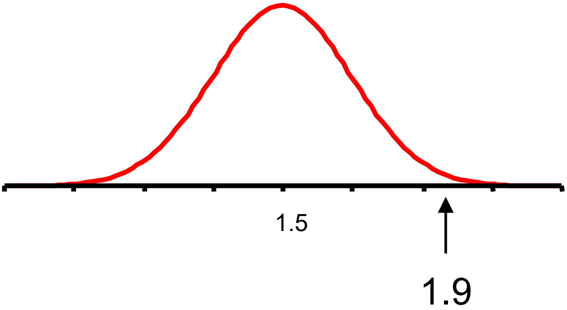
In this case, the sample mean ($\bar{x}$) is 1.9, the std dev (s) is .51 and n = 9

How far away is the sample mean from the required value of 1.5m?

$$t = \frac{\bar{x} - \mu}{\left(s / \sqrt{n} \right)} \sim T_{(n-1)}$$

$$= \frac{1.9 - 1.5}{.51 / \sqrt{9}}$$

$$= 2.35$$



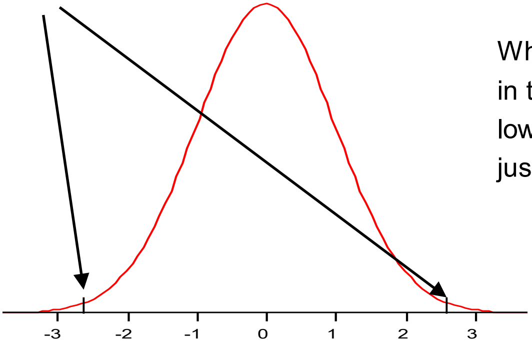
We now determine the probability of getting a sample mean this far away if the true mean really was 1.5m

Step 4 Determine the *p*-value

t-test statistic, $t = 2.35$ df=n-1=9-1=8

p-value = probability *t* is outside the range of *t* for both positive and negative values of *t*

ie $P(|t|> 2.35)$



Why do we look in the upper AND lower tails? Why not just one of them?

Step 5 Decision

p-value= 2 x .0233 = .0466 (using Excel 0.0465)
As p-value =.047 < 0.05 (significance level α),
reject the null hypothesis.

Step 6 Conclusion

Reject H_o and conclude that the data is NOT consistent with the claim that the average of growth of stinging leaf seedlings is 1.5m per year in direct sunlight.

Example 6: Whale mass

One sided *t*-test

Short-finned pilot whales had a mean weight of 360kg in the 1920s

Has the species been depleted so that the current mean weight is now less?

Step 1. The hypotheses

$$H_0 : \mu = 360kg$$

$$H_A : \mu < 360kg$$

Sample (n = 25):

Mean weight: $\bar{x}$= 336kg df=n-1=25-1=24

Standard deviation: s=30kg

Step 2. Test statistic

$$t = \frac{(336-360)}{30/\sqrt{25}} = (336-360)/6 = -4$$

Step 3. The sampling distribution of the test statistic

$$t = \frac{\bar{x}-\mu}{\left(s/\sqrt{n}\right)} \sim T_{(n-1)}$$

Notations:

- μ is the population mean under H0
- *s* is the sample standard deviation
- *n* is the sample size
- $\bar{x}$ is the sample mean

Step 4. p-value:

$P(t_{24} < -4) = 0.0003$ (look at HA)

Step 5. Decision:

As p-value is very small (say for a significance level α of 5%), that is p-value (0.0003)< α (0.05), then we reject H_o.

Step 6. Conclusion: Reject the null hypothesis and conclude that the mean whale weight is now less than 360kg

Assumptions

The following assumptions are made when conducting 1 sample tests.

1. The data is a random sample from the population of interest.
2. The distribution of sample means is Normal.

Which means either:

The underlying data distribution is Normal, or

The sample size is large enough for the CLT to give a Normal distribution for the sample means