Subject: Leaving Certificate Maths

Teacher: Mr Murphy

Lesson 14: Statistics III & Probability II

14.1 Learning Intentions

After this week's lesson you will be able to;

- Recall how z-scores work.
- Understand how to use sample proportion.
- Describe what is meant by standard error.
- Hypothesis Testing using (sample proportion).
- Describe what is meant by the central limit theorem.
- Outline what is meant by p-values.
- Calculate expected value
- Examine Bernoulli trials

14.2 Specification

Students learn about	Students working at OL should be able to	In addition, students working at HL should be able to
1.7 Analysing, interpreting and drawing inferences from data	– recognise how sampling variability influences the use of sample information to make statements about the population – use appropriate tools to describe variability drawing inferences about the population from the sample – interpret the analysis and relate the interpretation to the original question – interpret a histogram in terms of distribution of data – make decisions based on the empirical rule – recognise the concept of a hypothesis test – calculate the margin of error ($\frac{1}{\sqrt{n}}$) for a population proportion* – conduct a hypothesis test on a population proportion using the margin of error	– build on the concept of margin of error and understand that increased confidence level implies wider intervals – construct 95% confidence intervals for the population mean from a large sample and for the population proportion, in both cases using z tables – use sampling distributions as the basis for informal inference – perform univariate large sample tests of the population mean (two-tailed z-test only) – use and interpret p-values
	Numerical – recognise standard deviation and interquartile range as measures of variability – use a calculator to calculate standard deviation – find quartiles and the interquartile range – use the interquartile range appropriately when analysing data – recognise the existence of outliers	**Numerical** – recognise the effect of outliers – use percentiles to assign relative standing

14.3 Chief Examiner's Report

Section	Q	Mean Mark	Mean Mark (%)	Mark Ranking (Paper)	Main Topic
A	1	21·6	84	1	Probability
A	2	14·8	59	6	Inferential statistics

13.4 Z-Scores

Recall how to calculate probability from the z-score

Calculate $P(Z < 0.76)$ Read directly from pg 36/37

Calculate $P(Z > 1.76)$ Find 1 – prob that it's ≤ to the z-score

Calculate $P(Z \leq -0.76)$ Find 1 – prob that it's ≤ to the + z-score

Calculate $P(1.8 \leq Z \leq 2.1)$ Find P($Z \leq$ 2.1) – P($Z \leq 1.8$)

Recall $z - score = \dfrac{x-\mu}{\sigma}$ ……Pg 36

Question:

The mean height of footballers on the Irish national soccer
Team is 164cm with a standard deviation of 12cm. Assuming the heights are
normally distributed, find the % of players that are:

i) Less than 170cm in height

ii) Greater than 180 cm in height.

14.5 Hypothesis Testing

This is essentially testing whether a claim has some truth to it or not. We are testing a hypothesis. In a question, similar to week 13 you can be asked to deal with a proportion $(\hat{p})$ or mean $(\bar{x})$. In either case we set up our confidence intervals just like last week and work from there. Before we look at solving these problems, we need to understand two terms:

i) **Null Hypothesis**: Denoted by H_0 is a claim or statement about a population. We assume this statement is true until proven otherwise. (the null hypothesis means that nothing is wrong with the claim or statement).

ii) **Alternative Hypothesis**: Denoted by H_1 is a claim or statement which opposes the original statement about a population.

To carry out one of these tests we:

1 State the null hypothesis, H_0

2 State the alternative hypothesis, H_1

3 Find the confidence interval

$$\hat{p} - Error < p < \hat{p} + Error$$

or

$$\bar{x} - Error < \mu < \bar{x} + Error$$

(Calculate error with either $1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$) for proportions or $\frac{\sigma}{\sqrt{n}}$ for means

4 Reject or fail to reject the null hypothesis.

Question:

A clever student reckons he has devised an app to identify who has been viewing your SnapTok profiles. However, it only has a success rate of 80%.
The maker of SnapTok doubts this claim and carries out a test. In their test, 868 of 1100 times the app worked.

At the 95% level of confidence is this result in agreement with the student's claims?

Question:

A news website is reporting that the average video conference meeting lasts 34 minutes. In a random sample of 320 people working from home were surveyed. The mean time from the survey was 28 minutes with a standard deviation of 8 minutes.

Using a hypothesis test at the 5% level of significance, test the websites claim.

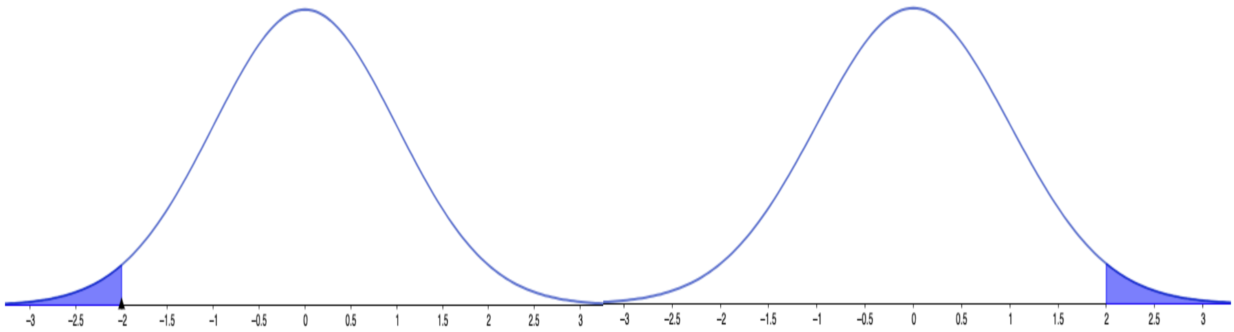
14.6 P-Values

P-values are another way of analysing data that may be asked of you in an exam. A p-value is a probability. What it measures is the probability that if you take a different sample group from your population, the mean of this sample group will be further away from the actual population mean.

To calculate a p-value we need our z-score and to refer to the normal distribution curve. You may see this formula when looking up p-values:

$$p - value = 2\left(1 - P(z \leq |T|)\right)$$

But let's look at where this formula is describing on the normal distribution.



The regions highlighted in blue are the areas where we have a result that is further away from the population mean that we would have looked at previously with our 95% confidence (z = 1.96). So, for a p-value we are looking to calculate the areas shaded in blue. (Refer to video for workings)

Question:

A car manufacturer claims that its maximum speed is 200mph, with a standard deviation of 4.8mph. As a test, a sample of 150 cars are tested to have a mean speed of 197mph. Using p-values at the 95% level of significance is their claim true?

14.7 Expected Value

This is over the lifetime of a game, what is the average outcome. In terms of games involving money, what is the average pay-out per player relative to the cost to play the game. It is important to remember that a game is not just a carnival type game.

In this context a game can be anything that involves making a bet. It is basically the average of what is going to happen.

To calculate the expected value, we multiply

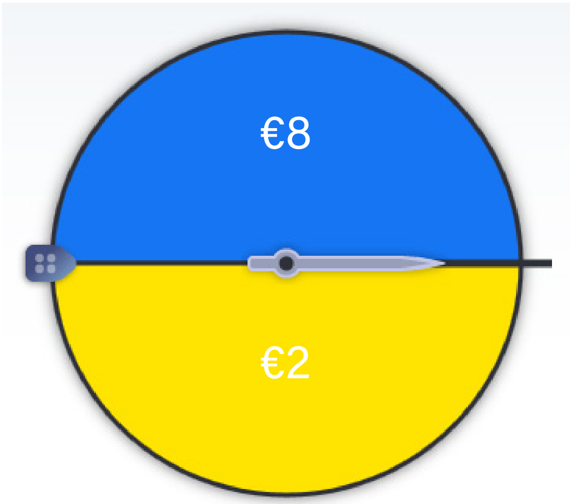
the value of the outcome

by

The probability of the outcome.

$$E(x) = \sum x.P(x)$$

Sample Game:



At a cost of €5 to play this game, would you play and why/why not?

Looking at the expected value we can decide if a game is biased or not. To give another real-world example of this, let's look at a popular gameshow, Deal or No Deal.

14.8 Bernoulli Trials

A Bernoulli trial is an event where there are two possible outcomes which we refer to as a success or a failure. Each outcome is independent of the other.

Question:

With 15 cars and 5 motorbikes in his garage, Conor McGreggor can choose anyone he likes to get to the gym from Monday – Friday.

Let's say we want to find the probability that he takes a car 3 days and a motorbike the other 2.

Copy down maths from video below:

In this event we refer to the number of trials, (days of the week we have to make a choice) as n.

r is the number of successes, 3 in this case.

Probability of success is p

Probability of failure is 1 – p a.k.a. q.

$$Probabilty = \binom{n}{r} p^r q^{n-r}$$

14.9 Recap of Learning Intentions

After this week's lesson you will be able to;

- Recall how z-scores work.
- Understand how to use sample proportion.
- Describe what is meant by standard error.
- Hypothesis Testing using (sample proportion).
- Describe what is meant by the central limit theorem.
- Outline what is meant by p-values.
- Calculuate expected value
- Examine Bernoulli trials

14.10 Homework Task

Question 1:

a. A motoring magazine collected data on cars on a particular stretch of road. Certain details on 800 cars were recorded.

   i. The ages of the 800 cars were recorded. 174 of them were new (less than 1 year old).
   Fint the 95% confidence interval for teh proportion of new cars on this road.
   Give your answer correct to 4 significant figures.

   ii. The data on teh speeds of these 800 vehicles is normally distributed with an average speed of 87.3 km per hour and a standard deviation of 12 km per hour.
   What proportion of cars on this stretch of road would you expect to find travelling at over 95 km per hours?

Question 2:

When Conor rings Ciara's house, the probability that Ciara answers the phone is . 1/5

a. Conor rings Ciara's house once every day for 7 consecutive days. Find the probability that she will answer the phone on the 2nd, 4th, and 6th days but not on the other days.

b. Find the probability that she will answer the phone for the 4^{th} time on the 7^{th} day.

c. Conor rings her house once every day for n days. Write, in terms of n, the probability that Ciara will answer the phone at least once.

14.11 Solutions to 13.10

Question: 60 marks

a. (i) Acme Confectionery makes cakes and chocolate bars. (a) (i) Acme Confectionery has launched a new bar called Chocolate Crunch. The weights of these new bars are normally distributed with a mean of 4.64 g and a standard deviation of 0.12 g. A sample of 10 bars is selected at random and the mean weight of the sample is found. Find the probability that the mean weight of the sample is between 4.6 g and 4 7 g.

b. (ii) A company surveyed 400 people, chosen from the population of people who had bought at least one Chocolate Crunch bar. Of those surveyed, 324 of them said they liked the new bar. Create the 95% confidence interval for the population proportion who liked the new bar. Give your answer correct to 2 decimal places.