FAMILY NAME: FIRST NAME:

ID No:

STATS 210 FC **Term Test** **Date: Thursday 14th April, 2016**

Answer **ALL THREE QUESTIONS**. Marks are shown for each question.

1. (a) Clara and Jake are looking for their mum's car key. It could be under the table, behind the bookshelf, or under the dishwasher, each with equal probability. Clara will only search under the table; if the key is there she has a $\frac{3}{5}$ chance of finding it. Jake will only search behind the bookcase; if the key is there he has a $\frac{1}{5}$ chance of finding it.

(i) Find the probability that the car key is found. (3)

Let T = "under table"; B = "behind bookshelf";
D = "under dishwasher"

$\mathbb{P}(T) = \mathbb{P}(B) = \mathbb{P}(D) = 1/3$

Let F = "found"

$\mathbb{P}(F|T) = 3/5$; $\mathbb{P}(F|B) = 1/5$; $\mathbb{P}(F|D) = 0$

$\mathbb{P}(F) = \mathbb{P}(F|T)\mathbb{P}(T) + \mathbb{P}(F|B)\mathbb{P}(B) + \mathbb{P}(F|D)\mathbb{P}(D)$

$= \frac{1}{3}\left(\frac{3}{5} + \frac{1}{5} + 0\right) = \frac{4}{15}$ $(=0.267)$

(ii) Suppose the car key has been found. Find the probability that it is found by Jake. (3)

Key can only be found by Jake if it is behind the bookcase.

So $\mathbb{P}(B|F) = \dfrac{\mathbb{P}(F|B)\mathbb{P}(B)}{\mathbb{P}(F)} = \dfrac{1/5 * 1/3}{4/15} = \frac{1}{4}$ $(=0.25)$

(iii) Suppose the car key has not been found. Find the probability that it is under the dishwasher. (3)

Looking for $\mathbb{P}(D|\bar{F})$:

$$\mathbb{P}(D|\bar{F}) = \frac{\mathbb{P}(\bar{F}|D)\,\mathbb{P}(D)}{\mathbb{P}(\bar{F})}$$

$$= \frac{1 * 1/3}{1 - 4/15} = \frac{1/3}{11/15} = \frac{5}{11} \quad (= 0.455)$$

(b) Out of all students graduating with a first class honours degree in statistics from a particular university, 94% passed the mid-semester test in the *Statistical Theory* course. Jim concludes that since he has just passed the mid-semester test, he has an excellent chance of getting a first class honours degree.

(i) Explain the main flaw in this argument. (3)

Let F = "first class honours", T = "Pass Stat Theory mid-semester test"

We are told that $\mathbb{P}(T|F) = 0.94$

Jim is assessing the probability of getting a first class honours degree given that he has passed the test.

This is the conditional event $\mathbb{P}(F|T)$.

The flaw in Jim's argument is that he is confusing $\mathbb{P}(T|F)$ with $\mathbb{P}(F|T)$, which are not the same probabilities.

(ii) Suppose that 80% of students pass the test and 10% of students get a first class honours degree. Find the probability that a randomly selected student who passed the test gets a first class honours degree. (4)

Information given:

$\mathbb{P}(F) = 0.1$; $\mathbb{P}(T) = 0.8$; $\mathbb{P}(T|F) = 0.94$ (from part(i))

We are looking for $\mathbb{P}(F|T)$:

$$\mathbb{P}(F|T) = \frac{\mathbb{P}(T|F)\,\mathbb{P}(F)}{\mathbb{P}(T)} = \frac{0.94 * 0.1}{0.8}$$

$$= \frac{0.094}{0.8} = \underline{\underline{0.1175}}$$

2. The **Fortune 500 list** is a list of the 500 richest businesses in the US. In 2005, a study was done of the chief executive officers (CEOs) of 250 businesses on this list, all of whom were male. It was discovered that 58% of the CEOs were over 6 feet tall (183cm). In the US population, 14.5% of adult men are over 6 feet tall. We wish to test whether these figures provide evidence that there are more tall men among US company CEOs than we would expect by chance.

(a) Let X be the number of CEOs over 6 feet tall from the 250 businesses studied. Formulate the null hypothesis, H_0, and the alternative hypothesis, H_1, in terms of the distribution of X and its parameters. Remember to specify the full distribution of X and to use a two-sided alternative hypothesis. (3)

If CEOs were representative of the US population with respect to height, we would expect 14.5% of them to be taller than 6 feet (183cm).

Let X = # of the 250 surveyed CEOs > 6 feet tall.

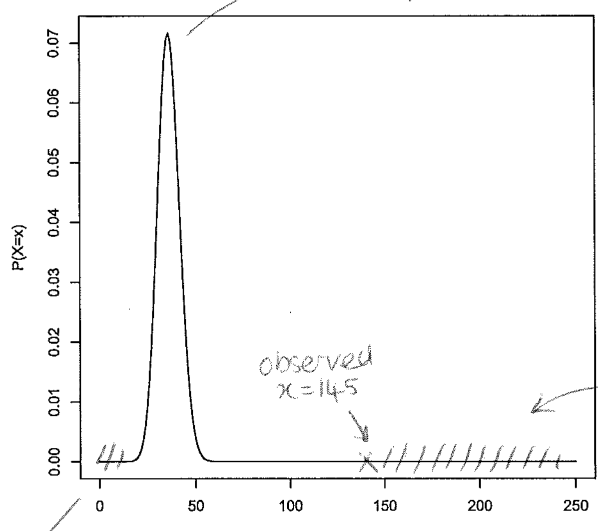
$X \sim \text{Binomial}(250, p)$

$H_0: p = 0.145$

$H_1: p \neq 0.145$

(b) The probability function of X under the null hypothesis is shown below. Find the approximate value of x where the curve peaks. Also mark the observed value of x so that you can see the tail probabilities required for the p-value, and shade under the curve the area represented by the p-value. (3)

curve peaks at 250×0.145 $= 36.25$ under H_0.



Observed value $= 0.58 \times 250 = 145$

$\mathbb{P}(X \geq 145)$

equal prob in lower tail

(c) Estimate the p-value for the hypothesis test using the graph of the probability function above. Interpret the result in terms of the strength of evidence against the null hypothesis. Is the observed data compatible with the null hypothesis? (4)

p-value = $2 * \mathbb{P}(X \geq 145)$ where $X \sim \text{Binomial}(250, 0.145)$

Judging by the graph of the probability function of X, the p-value is extremely small, probably indistinguishable from zero.

We therefore have extremely strong evidence against H_0, i.e. extremely strong evidence that these CEOs were not drawn at random from the US population. There is virtually no chance that a random sample of 250 US men would yield as many as 145 over 6 feet tall purely by chance.

3. In football, a penalty kick is awarded when a foul that is punishable by a direct free kick is committed within the offending player's own penalty area. Penalties are converted into goals (called a *conversion*) more often than not, even against very talented goal keepers. In the English Premier League 2014-2015 season, a total of 83 penalties were awarded. The conversion was successful on 63 of the 83 opportunities*.

*Note: data available at:
`http://www.myfootballfacts.com/Premier_League_Penalty_Statistics.html`.

Suppose that every 2014-2015 season English Premier League conversion is successful with probability p, independently of all other conversions. Our data are 63 successful conversions out of 83 attempts. We wish to estimate the probability p that any given conversion is successful.

(a) Write down the likelihood function, $L(p; x)$, substituting the correct value of x. State the range of values of p for which the likelihood function is defined. (2)

$$L(p;63) = \binom{83}{63} p^{63} (1-p)^{83-63}$$

$$= \binom{83}{63} p^{63} (1-p)^{20} \quad \text{for } 0 < p < 1.$$

(b) Find $\frac{dL}{dp}$, and give all possible solutions to the equation $\frac{dL}{dp}=0$. (4)

$$\frac{dL}{dp} = \binom{83}{63}\left\{63p^{62}(1-p)^{20} - 20p^{63}(1-p)^{19}\right\}$$

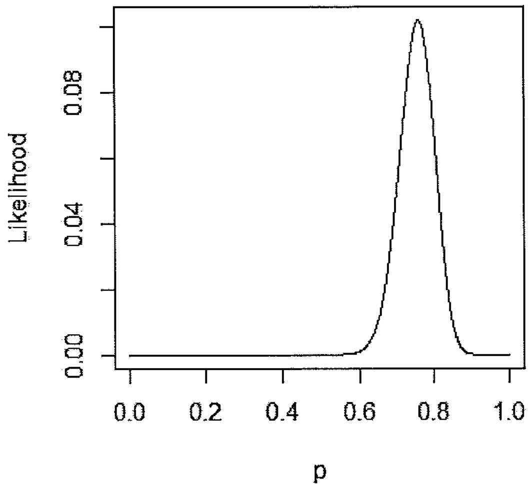
$$= \binom{83}{63} p^{62}(1-p)^{19} \left\{63(1-p) - 20p\right\}$$

$$= \binom{83}{63} p^{62}(1-p)^{19} (63-83p)$$

$$\left.\frac{dL}{dp}\right|_{p=\hat{p}} = 0 \quad \text{when } p=0,\ p=1,\ p=\frac{63}{83}.$$

Possible solutions: $p=0$, $p=1$, $p=63/83$.

(c) The likelihood function is plotted below. By referring to the graph and using your answer for (b), find the maximum likelihood estimate of p and state what this maximum likelihood value represents. (3)



It is clear from the graph of the likelihood that the maximum does not occur at $p=0$ or $p=1$.

The maximising value is $p = 63/83$.

Thus the MLE is $\hat{p} = 63/83$.

The MLE of p, $\hat{p}$, is the value of p for which the observation of $x=63$ is more likely than at any other value of p.

End of Paper **Total: 35**