Question 17

Let the random variable $\hat{P}$ represent a sample proportion observed in an experiment.

If $p = \frac{1}{6}$, what is the smallest integer value of the sample size such that the standard deviation of $\hat{P}$ is less than or equal to $\frac{1}{216}$?

A. 30

B. 6480

C. 180

D. 1600

E. 16

$$\sigma(\hat{p}) = \sqrt{\frac{\frac{1}{6} \times \frac{5}{6}}{n}}$$

$$\therefore \sqrt{\frac{5}{36n}} \leq \frac{1}{216}$$

$$\therefore n \geq 6480$$

Question 18

An exit poll of 1000 voters found that 620 favoured candidate A.

An approximate 90% confidence interval for the proportion of voters from the total population in favour of candidate A is

A. $\left(0.62 - \sqrt{\frac{0.62 \times 0.38}{1000}},\ 0.62 + \sqrt{\frac{0.62 \times 0.38}{1000}}\right)$

B. $\left(0.62 - 1.65\sqrt{\frac{0.62 \times 0.38}{1000}},\ 0.62 + 1.65\sqrt{\frac{0.62 \times 0.38}{1000}}\right)$

C. $\left(0.62 - 2.58\sqrt{\frac{0.62 \times 0.38}{1000}},\ 0.62 + 2.58\sqrt{\frac{0.62 \times 0.38}{1000}}\right)$

D. $\left(620 - 1.96\sqrt{\frac{0.62 \times 0.38}{1000}},\ 620 + 1.96\sqrt{\frac{0.62 \times 0.38}{1000}}\right)$

E. $\left(0.62 - \sqrt{\frac{0.62 \times 0.38}{620}},\ 0.62 + \sqrt{\frac{0.62 \times 0.38}{620}}\right)$

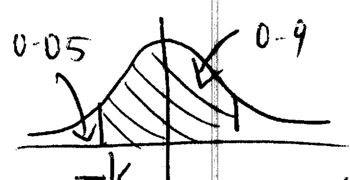
$$\hat{p} = \frac{62}{1000} = 0.62$$

Interval:

$$\left(\hat{p} \pm k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

$\hat{p} = 0.62$

$n = 1000$



$-k = \text{invNorm}(0, 1, 0.05)$

$-k \approx -1.645$

Confidence interval: $\left(0.62 - 1.645\sqrt{\frac{0.62 \times 0.38}{1000}},\ 0.62 + 1.645\sqrt{\frac{0.62 \times 0.38}{1000}}\right)$

A randomly selected group of 820 people on the electoral role were asked whether they favoured fixed five year terms for Federal parliament. Fifty-three percent of the group favoured the idea. An approximate 95% confidence interval for the proportion p of people on the electoral role who favour the idea can be found by calculating

A. $0.53-\sqrt{\dfrac{0.53\times 0.47}{820}}<p<0.53+\sqrt{\dfrac{0.53\times 0.47}{820}}$

B. $0.53-0.95\sqrt{\dfrac{0.53\times 0.47}{820}}<p<0.53+0.95\sqrt{\dfrac{0.53\times 0.47}{820}}$

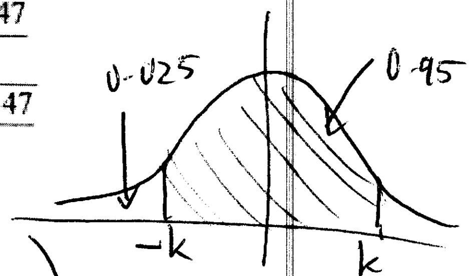
C. $0.53-1.64\sqrt{\dfrac{0.53\times 0.47}{820}}<p<0.53+1.64\sqrt{\dfrac{0.53\times 0.47}{820}}$

D. $0.53-1.96\sqrt{\dfrac{0.53\times 0.47}{820}}<p<0.53+1.96\sqrt{\dfrac{0.53\times 0.47}{820}}$

E. $0.53-2.58\sqrt{\dfrac{0.53\times 0.47}{820}}<p<0.53+2.58\sqrt{\dfrac{0.53\times 0.47}{820}}$

$\hat{p}=0.53$

$\left(0.53-1.96\sqrt{\dfrac{0.53\times 0.47}{820}},\ 0.53+1.96\sqrt{\dfrac{0.53\times 0.47}{820}}\right)$



$-k = \text{invNorm}(0.025, 0, 1)$

$-k = -1.96$

Question 17

In a population of tropical fish, 15% have a disease. A random sample of 300 of the fish is taken. A normal approximation is used to find the approximate probability that less than 20% of the fish in the sample have the disease. That approximate probability is closest to

A. 0.97938
B. 0.98326
C. 0.98723
D. 0.99225
E. 0.99235

$X \overset{d}{=} Bi(n=300, p=0.15)$

$\mu = np = 300\times 0.15 = 45$

$\sigma = \sqrt{npq} = \sqrt{300\times 0.15\times 0.85} \approx 6.1846584$

$X \overset{d}{\simeq} N(\mu = 45, \sigma = 6.1846584)$

If $\hat{p} = 20\%$, $X = 0.2\times 300 = 60$

$Pr(X \leq 60) \simeq 0.99235$

Question 3 (17 marks)

A bank has over one million customers. The proportion of these customers who don't use online banking is $\frac{2}{9}$.

The bank takes a number of random samples of its customers. Each sample contains 15 customers. Let X be the random variable that represents the number of customers in a sample who don't use online banking.

a. Find $\Pr(X \le 5)$. Give your answer correct to three decimal places.

2 marks

$X \stackrel{d}{=} Bi\left(n = 15, p = \frac{2}{9}\right)$

$\Pr(X \le 5) = 0.906$

Let $\hat{P}$ be the random variable of the distribution of sample proportions of customers who don't use online banking.

b. Find

i. the expected value of $\hat{P}$

.

1 mark

$E(\hat{p}) = p = \frac{2}{9}$

ii. the standard deviation of $\hat{P}$

.

2 marks

$\sigma(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$

$= \sqrt{\frac{\frac{2}{9} \times \frac{7}{9}}{15}} = \frac{\sqrt{210}}{135}$

(Exact value required!)

c. Suppose the bank wanted a standard deviation of $\hat{P}$ of less than 0.1.

What is the minimum number of customers that the bank would need to have in each sample in order to achieve
this? 2 marks

$\sigma < 0.1$

$\therefore \sqrt{\dfrac{\frac{2}{9} \times \frac{7}{9}}{n}} < 0.1$

$\therefore n > 17.28$

$\therefore n = 18$ would be the minimum.

d. The bank decides to retain its original sample size of 15 customers.
Find the probability that a sample proportion would lie within one standard deviation of the population proportion. Do not use a normal approximation. Give your answer correct to three decimal
places.

3 marks

$$\Pr(p - \sigma < \hat{p} < p + \sigma)$$

$$= \Pr\left(\frac{2}{9} - \frac{\sqrt{210}}{135} < \hat{p} < \frac{2}{9} + \frac{\sqrt{210}}{135}\right)$$

$$= \Pr(0.114879 < \hat{p} < 0.329566)$$

But $\hat{p} = \frac{X}{15}$ $\therefore X = 15\hat{p}$

$$= \Pr(1.72 < X < 4.94)$$

$$= \Pr(2 \leq X \leq 4)$$

Where $X \stackrel{d}{=} Bi\left(n=15, p=\frac{2}{9}\right)$

$\Pr(2 \leq X \leq 4) \simeq$ 0.652

Tom also has parsley plants in his nursery. He obtains his parsley plants from growers all over Australia. It is known that 30% will be prone to a particular leaf disease.

Tom decides to test his plants for the leaf disease. He takes a random sample of 20 parsley plants.

g. **i.** What is the probability that the sample proportion is equal to the population proportion of 0.3? Give your answer correct to four decimal places. Do not use a normal approximation.

2 marks

$\hat{p} = \frac{X}{20}$

$\Pr(\hat{p} = 0.3) = \Pr(X = 0.3 \times 20)$

$= \Pr(X = 6)$

where $X \stackrel{d}{=} Bi(n = 20, p = 0.3)$

$\Pr(X=6) = 0.1916$

ii. What is the probability that the sample proportion lies within two standard deviations of the population proportion? Give your answer correct to four decimal places. Do not use a normal approximation.

3 marks

$\Pr(p - 2\sigma < \hat{p} < p + 2\sigma)$

$\sigma = \sqrt{\frac{p(1-p)}{20}} = \sqrt{\frac{0.3 \times 0.7}{20}} = \frac{\sqrt{105}}{100}$

$\Pr\left(0.3 - 2 \times \frac{\sqrt{105}}{100} < \hat{p} < 0.3 + 2 \times \frac{\sqrt{105}}{100}\right)$

$= \Pr(0.09506098 < \hat{p} < 0.50493902)$

$= \Pr(1.9 < X < 10.098)$

$= \Pr(2 \le X \le 10)$ where $X \stackrel{d}{=} Bi(n = 20, p = 0.3)$

$= 0.9752$

While testing for the leaf disease, Tom sees that some of his plants are infested with a leaf moth. He finds that 236 plants out of 500 are infested with the leaf moth.

h. Find an approximate 95% confidence interval for the proportion of plants infested with the leaf moth. Give your answer correct to three decimal places.

1 mark

$\hat{p} = \frac{236}{500}$ $X = 236$, $n = 500$, $CL = 0.95$

Interval: $(0.428, 0.516)$

Another 26 nursery owners from around Australia independently sample parsley plants from their stocks and test for leaf disease. Each calculates an approximate 95% confidence interval for p, the proportion of plants in the population infested with the leaf moth. It is subsequently found that of these 27 confidence intervals exactly one does not contain the value of p.

Researchers investigating the prevalence of leaf moth randomly select five of the confidence intervals calculated by the nursery owners.

i. What is the probability that exactly three of the selected confidence intervals contain the value of p? Give your answer correct to four decimal places.

2 marks

Let X = no. of confidence intervals that contain the value of p

$X \stackrel{d}{=} Bi\left(n = 5, p = \frac{26}{27}\right)$

$Pr(X = 3) = 0.0122$

Question 1

The frequency of colour blindness in the Caucasian male American population is estimated to be about 8%. We take a random sample of 25 from this population.

a. Calculate the probability that more than 3 in the sample will be colour blind?

$X \stackrel{d}{=} Bi(n=25, p=0.08)$ $X =$ no. who are colour blind

$X = 0, 1, 2, \ldots, 25$

$Pr(X>3) = Pr(X \geq 4) = 0.1351$

b. Calculate the probability that exactly no more than 1 in the sample is colour blind.

$Pr(X \leq 1) = 0.3947$

c. Why would a normal approximation to the binomial distribution not be valid in this context?

$np = 25 \times 0.08 = 2 < 5$

Binomial distribution is too skewed for Normal approximation to be justified.

d. If a random sample of size 125 is instead selected, explain why the Normal approximation to the Binomial distribution would now be justified.

$np = 125 \times 0.08 = 10 > 5$

Now, with this larger sample size, the Binomial distribution is sufficiently symmetrical to justify the Normal approximation.

e. In a random sample of 125, it was found that 7 males had colourblindness.

i. What is the value of the sample proportion, $\hat{p}$?

$\hat{p} = \frac{7}{125} = 0.056$

ii. Assuming that the population proportion is 8%, does the value of $\hat{p}$ calculate above lie within a 95% confidence interval for the population proportion?

95% Confidence interval based on $p = 0.08$

$\therefore X = 0.08 \times 125 = 10$

$n = 125$

$CL = 0.95$ gives: $(0.0324, 0.1276)$

Yes, $\hat{p} = 0.056$ lies in this interval.

Question 2

In a sample of 519 judges, it was found that 285 were introverts. Let p = the proportion of all judges who are introverts.

a. Find a point estimate for p.

Point estimate $= \hat{p} \simeq \frac{285}{519} \simeq 0.5491$

b. Find a 99% confidence interval for p and explain what it means.

$X = 285$

$n = 519$

$CL = 0.99$ gives $(0.4929, 0.6054)$

There is a probability of 0.99 that the true overall proportion of introverted judges lies within the interval $(0.4929, 0.6054)$

Question 3

A survey asked a nation wide random sample of 2500 adults if they agreed or disagreed with the statement: " I like buying new clothes, but shopping is often frustrating and time-consuming". Suppose that 60% of all Australian adults would agree with this statement. Calculate the probability that at least 1520 people in the sample agree,

i. not using the Normal approximation.

$X \stackrel{d}{=} Bi(n=2500, p=0.6)$ $\quad$ X = no. of adults who agree

$Pr(X \geq 1520) \simeq 0.2131$

ii. Using the Normal approximation.

$X \stackrel{d}{\simeq} N(\mu=1500, \sigma=24.494897)$

$\mu = np$
$= 2500 \times 0.6$
$= 1500$

$\sigma = \sqrt{2500 \times 0.6 \times 0.4}$
$= 24.494897$

$Pr(X \geq 1520) \simeq 0.2071$

Exercise 5

Casey buys a Venus chocolate bar every day for 180 days, during a promotion promising that 'one is six wrappers is a winner'. From these 180 purchases, Casey gets 20 winning wrappers.

a What is the proportion of winning wrappers Casey expects to get, if the advertised claim is true? How many winning wrappers would this imply, for Casey?

b What is the proportion of winning wrappers in Casey's sample?

c Find an approximate 95% confidence interval for the true proportion of winning wrappers, based on Casey's sample of Venus bars.

d Casey feels he has missed out, and suspects that the true proportion of winners is not one in six. Comment on this, based on Casey's sample of Venus bars.

e What assumptions have been made about Casey's sample of Venus bar wrappers?

(a) $\frac{1}{6}$

(b) $\hat{p} = \frac{20}{180} = \frac{1}{9}$

(c) $X = 20$

$n = 180$

$CL = 0.95$ gives $(0.0652, 0.1570)$

(d) There is a probability of 0.95 that this interval contains the true proportion of winners.

Since $1/6 \approx 0.167$ is not in this interval, it is 95% probable that $\frac{1}{6}$ is not the true proportion of winners.

(e) That Casey's sample is random (so, for example, that he has randomly selected the places where he purchased Venus bars)

Also that $180p > 5$ (where p is the true proportion of winners in the population) so that the normal approximation applies.