

i Cover Page

Cover Page

Department of Economics

Examination paper for SØK1004 Statistics for Economists

Examination date: 21.12.2022

Examination time (from-to): 15:00 – 19:00

Permitted examination support material: A / All support material is allowed

The support material is also attached to the exam set (you can find it at the bottom of the page throughout the exam).

Academic contact during examination: Quynh Trang Nguyen

Phone: 90850838

Technical support during examination: Orakel support services

Phone: 73 59 16 00

If you experience technical problems during the exam, contact Orakel support services as soon as possible before the examination time expires/the test closes. If you don't get through immediately, hold the line until your call is answered.

OTHER INFORMATION

Do not open Inspira in multiple tabs, or log in on multiple devices, simultaneously. This may lead to errors in saving/submitting your answer.

Read the questions carefully and make your own assumptions. If a question is unclear/vague, make your own assumptions. Only contact academic contact in case of errors or insufficiencies in the question set.

Cheating/Plagiarism: The exam is an individual, independent work. During the exam it is not permitted to communicate with others about the exam questions or distribute drafts for solutions. Such communication is regarded as cheating. All submitted answers will be subject to plagiarism control.

Notifications: If there is a need to send a message to the candidates during the exam (e.g. if there is an error in the question set), this will be done by sending a notification in Inspira. A dialogue box will appear. You can re-read the notification by clicking the bell icon in the top right-hand corner of the screen.

ABOUT SUBMISSION

Answering in Inspira: If the question set contains questions that are not upload assignment, you must answer them directly in Inspira. In Inspira, your answers are saved automatically every 15 seconds.

NB! We advise against pasting content from other programs, as this may cause loss of formatting and/or entire elements (e.g. images, tables).

Automatic submission: Your answer will be submitted automatically when the examination time expires and the test closes, as long as you have answered at least one question. This will happen even if you do not click “Submit and return to dashboard” on the last page of the question set. You can reopen and edit your answer as long as the test is open. If no questions are answered by the time the examination time expires, your answer will not be submitted. This is considered as “did not attend the exam”.

Withdrawing from the exam: If you become ill or wish to submit a blank test/withdraw from the exam for another reason, go to the menu in the top right-hand corner and click “Submit blank”. This cannot be undone, even if the test is still open.

Access to your answers: You will find your answer in Archive when the examination time has expired.

1 New Question: 1 point

A graphical presentation of a frequency distribution, relative frequency distribution, or percent frequency distribution of quantitative data constructed by placing the class intervals on the horizontal axis and the frequencies on the vertical axis is a _____.

Select one alternative:

- scatter plot
- bar chart
- histogram
- pie chart

Maximum marks: 1

2 New Question: 1 point

The sample mean's sampling distribution approaches a _____ when the number of observations increases.

Select one alternative:

- binomial distribution
- Poisson distribution
- normal probability distribution
- hypergeometric distribution

Maximum marks: 1

3 New Question: 1 point

Your friend claims that his secret brownie recipe satisfies at least 90% of the people whom he has given the brownies to. You plan on taking a sample to test his claim. The correct set of hypotheses is _____.

Select one alternative:

- $H_0: p \geq 90$ $H_a: p < 90$
- $H_0: p \leq 0.9$ $H_a: p > 0.9$
- $H_0: p \leq 90$ $H_a: p > 90$
- $H_0: p \geq 0.9$ $H_a: p < 0.9$

Maximum marks: 1

4 New Question: 1 point

A Type I error is committed when _____.

Select one alternative:

- a false alternative hypothesis is accepted *(i.e., you reject the true null hypothesis)*
- sample data contradict the null hypothesis
- a false null hypothesis is accepted
- the critical value is greater than the value of the test statistic

Maximum marks: 1

5 New Question: 1 point

From a normal population, a random sample of 100 observations is taken. The sample standard deviation is 23.

23 is a _____.

Select one alternative:

- point estimate
- point estimator
- population parameter because the sample is big enough
- sample parameter because it is calculated from the sample

Maximum marks: 1

6 **New Question: 1 point**

The t distribution gets closer to the _____ distribution when the sample size increases.

Select one alternative:

- uniform
- F
- Chi squared
- normal

Maximum marks: 1

7 **New Question: 1 point**

When the sample size decreases, the _____.

Select one alternative:

- standard deviation of the population increases
- standard deviation of the sample mean increases
- population mean decreases
- standard deviation of the sample mean decreases

$$s^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

As $n \downarrow \rightarrow s^2 \uparrow \rightarrow \sqrt{s^2} \uparrow$

Maximum marks: 1

8 **New Question: 1 point**

The sampling distribution of the ratio of two independent sample variances taken from normal populations with equal variances is a(n) _____ distribution.

Select one alternative:

- F
- normal
- Chi squared
- t

Maximum marks: 1

9 **New Question: 1 point**

Reduce the sample size by half will _____.

Select one alternative:

- reduce the standard error of the mean to approximately 0.71 times of its current value
- have no effect on the standard error of the mean
- increase the standard error of the mean to 1.41 times of its current value
- increase the standard error of the mean to double its current value

$$S.E_1 = \frac{\sigma}{\sqrt{n}} \quad S.E_2 = \frac{\sigma}{\sqrt{\frac{1}{2}n}} = \frac{1}{\sqrt{\frac{1}{2}}} \cdot \frac{\sigma}{\sqrt{n}} = 1.41(S.E_1)$$

Maximum marks: 1

10 New Question: 1 point

What is the difference between a bar graph and a histogram?

Select one alternative:

- A bar graph has no spaces between the bars, while a histogram must have space between the bars.
- A bar graph displays categorical data, while a histogram displays quantitative data.
- A bar graph displays quantitative data, while a histogram displays categorical data.
- There is no difference between a bar graph and a histogram.

Maximum marks: 1

11 New Question: 1 point

Less evidence against H_0 (the null hypothesis) is indicated by _____.

Select one alternative:

- larger p-values
- higher levels of significance
- larger critical values
- higher probabilities of a Type II error

Smaller p-value \rightarrow More evidence against H_0

Maximum marks: 1

12 **New Question: 1 point**

Assume X is a continuous random variable for $0 < X < 1$, $P(X=0.5)$ is _____.

Select one alternative:

0

0.5

1

unknown because we don't know the distribution of X

When X is a continuous RV
 $\rightarrow P(X=k) = 0$ for all k .
 We only work with intervals!

Maximum marks: 1

13 **New Question: 1 point**

During the corona lockdown, you started a new hobby: sourdough baking. Within seven days, you tried different temperatures to store the dough in before baking: room temperature (22°C), fridge temperature (2°C), and freezer temperature (-18°C). The variance of this seven-day-temperature trial is _____.

Select one alternative:

400

negative

266.67

non-negative

You do not have enough info to calculate the variance!
 But you know the variance is the square of the standard deviation \Rightarrow always ≥ 0 .

Maximum marks: 1

14 **New Question: 1 point**

An experiment is conducted where one of the outcomes is A with $P(A) = \frac{1}{5}$. You performed the experiment four times and A did not happen. On the fifth trial, A _____.

Select one alternative:

- must occur
- will not occur
- may occur
- has an 80% of occurring

$$P(A) = \frac{1}{5} \text{ for each time}$$

Maximum marks: 1

15 **New Question: 1 point**

The health department of a town is considering a plan to increase the proportion of people taking the flu shots, which is currently around 0.07. The alternative hypothesis of the correct set of hypotheses for testing the plan's effect is

Select one alternative:

- $p \geq 0.07$
- $p > 0.07\%$
- $p > 0.07$
- $p \geq 0.07\%$

$$H_0: p \leq 0.07$$

$$H_a: p > 0.07 \rightarrow \text{research question}$$

Maximum marks: 1

16 New Question: 2 points

```
In [1]: import pandas as pd
data = 'Downloads\IR.xlsx'
data = pd.read_excel(data)
display(data)
```

→ WRONG!

You have downloaded an Excel file called *IR* in the Downloads folder on your computer to do your Statistics homework. You then open a new Jupyter notebook in a separate Statistics folder and read the file as in the attached snapshot.

Is this statement True or False? "These lines of code will help you read the excel file"

Select one alternative:

True

False

The relative path that you specified is wrong!

Maximum marks: 2

17 New Question: 2 points

```
In [1]: df = pd.DataFrame(np.random.rand(5,3),
                    columns=['Cashew', 'Hazelnut', 'Almond'])
df.plot(kind='bar', grid=False)
df.T.plot(kind='bar', stacked=True)
```

→ Lack: `import numpy as np`
`import pandas as pd`

You open a new Jupyter notebook and you would like to draw some bar graphs. The first cell of your notebook is shown in the attached snapshot.

Is this statement True or False? "The above lines of code will give you a bar chart and a stacked bar chart"

Select one alternative:

True

False

You have not imported the appropriate packages (numpy and pandas) to your notebook for Python to recognize the abbreviated names such as 'pd' or 'np'.

Maximum marks: 2

18 New Question: 1 point

A pet store has 3000 kinds of sales items. You work part-time at that store and you have a boring day at work. Since you are very interested in the material of your statistics course, you take a random sample 100 items from the store and find that 60% are items for dogs. You then conclude that 60% of all sales items at this store are for dogs.

What you have done here is *statistical inference*. True or False?

Select one alternative:

True

False

Maximum marks: 1

19 New Question: 1 point

~~Is this statement true or false?~~ You work with a two-tailed test and calculate the test statistic. The test statistic is negative (on the left hand side of the normal curve).

Is this statement True or False? "*The area under the curve to the left hand side of your test statistic gives you the p-value*"

Select one alternative:

True

False

You have to multiply this area by 2 to get the p-value for this two-tailed test.

Maximum marks: 1

20 **New Question: 1 point**

≠ 22

20,000 students in a university has a mean age of 22, a median of 21 and a mode of 21. The sample standard deviation is 2.5.

Is this statement True or False? "The age of the students at this university follow a normal distribution."

Select one alternative:

False

For a normal distribution, mean = median = mode.

True

Maximum marks: 1

21 **New Question: 1 point**

Over time, the average proportion of undergraduate students is 80%, and 20% are postgraduate students. The university wants to attract more postgraduate students, so they try a new marketing strategy for their postgraduate programs. A year after the new strategy, 60 out of the total 250 students in a faculty take a postgraduate program (which previously had the same 80:20 ratio between undergraduate and postgraduate students). The dean of the faculty would like to test whether the new strategy worked for the faculty.

Is this statement True or False? "This is an upper-tail test".

Select one alternative:

False

True

$$H_0 : p \leq 0.2$$

$$H_a : p > 0.2 \rightarrow \text{research question}$$

\rightarrow upper-tail test

Maximum marks: 1

22 New Question: 2 points

You play a game that can simulate different scenarios for Lord of the Rings, which will result in either Merry or Pippin reaching Mount Doom to destroy the Ring. You play the Pippin scenario 31 times, which shows a sample average of 21 months and 18 days and a variance of 28 days to reach Mount Doom. For the 26 times you play the Merry scenario, the average time to reach Mount Doom is 15 months and 19 days and the variance is 14 days. (There is no guarantee in any scenario that the Ring actually gets destroyed.)

Is this statement True or False? "You are testing the time variances of the two kinds of scenarios. Given a significance level of 0.1, the critical value is found to be 1.92"

Select one alternative:

True

False

whether there is a signif. difference between
Pippin has variance > Merry → Pippin
is population 1
 $\Rightarrow n_1 = 31, s_1^2 = 28$
 $n_2 = 26, s_2^2 = 14$
 $\alpha = 0.1$

$$F = \frac{s_1^2}{s_2^2} = \frac{28}{14} = 2 \left\{ \begin{array}{l} \rightarrow F_{\alpha/2, \text{num. df}} \\ \text{den. df} \end{array} \right.$$

num. df = $n_1 - 1 = 30$
den. df = $n_2 - 1 = 25$

$F_{0.05} = 1.92$

Two-tailed

Maximum marks: 2

23 New Question: 1 point

Credit card number is an example of categorical data.

Select one alternative:

False

True

Maximum marks: 1

24 New Question: 1 point

You don't need to measure the fairness of a coin (the probability of getting heads or tails) to compute a binomial probability problem regarding that coin.

Select one alternative:

- False You should not assume $p = P(\text{Heads}) = 0.5$ all the time!
 $P(\text{Tails}) = 0.5$

True

It could be the case that $p = 0.6$ or $p = 0.3$, etc....

Maximum marks: 1

25 New Question: 2 points

You get to travel to a fantasy world and bring back anything you like. You decide to enter the world created by George R. R. Martin and visit the Targaryens, who claimed that their blood purity was 100%. You sneak around (safely) and manage to get 16 blood samples from 16 different Targaryens. You are now back to the real world and ready to test the Targaryens' claim. From this hair sample, you get an average blood purity of 94% with a standard deviation of 8%.

Is this statement True or False? "The 99% confidence interval estimate of the Targaryen blood purity has an upper limit value of 99%"

Select one alternative:

False

True

σ unknown

$$\bar{X} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$\Rightarrow 0.94 \pm 2.947 \left(\frac{0.08}{\sqrt{16}} \right)$$

$$0.94 \pm 0.05894$$

$$\rightarrow (0.88106; 0.99894) \Rightarrow \text{Upper limit} = 99.89\% \approx 100\%$$

$$\bar{X} = 0.94$$

$$s = 0.08$$

$$n = 16 \Rightarrow df = 15$$

$$\alpha = 0.01 \Rightarrow \frac{\alpha}{2} = 0.005$$

$$\rightarrow t_{0.005, 15} = 2.947$$

Maximum marks: 2

26 New Question: 2 points

You get to travel to a fantasy world and bring back anything you like. You decide to enter the world created by George R. R. Martin and visit the Targaryens, who claimed that their blood was 100% pure. You sneak around (safely) and manage to get 16 blood samples from 16 different Targaryens. You are now back to the real world and ready to test the Targaryens' claim. From this hair sample, you get an average blood purity of 94% with a standard deviation of 8%.

Is this statement True or False? "With a significance level of 0.01, the p-value for this hypothesis test must be smaller than 0.005."

Select one alternative:

False

True

$$H_0: \mu = 100\% \quad (\text{the average blood purity})$$

$$H_a: \mu \neq 100\%$$

$$\text{Test stat: } t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{0.94 - 1}{0.08/\sqrt{16}} = -3 \rightarrow t \text{ upper tail} = 3$$

$$df = 15 \rightarrow p \text{ upper tail} < 0.005$$

Maximum marks: 2

$\downarrow \times 2$
 $\rightarrow p\text{-value (two-tailed)} < 0.01$

27 New Question: 1 point

We gain more evidence against the null hypothesis when the p-value becomes smaller.

Select one alternative:

False

True

Maximum marks: 1

28 New Question: 2 points

Top 100 movies rated by regular IMDb voters (as of 25. November, 2022) shows movies with *the highest* ratings. The standard deviation of the ratings is 0.2.

Your friend takes a random sample of 10 movies from this list and says that the average rating of the movies on this list is 8.5. Since you have taken the Statistics course and you tend to correct people, you tell your friend with 95% confidence that his estimate has a margin of error of

± 0.12

(Write the answer in TWO decimal places and separate the ones

and tenths by ".", not ",", Eg: 2.32, 3.54, etc.)

$$\text{known } \Rightarrow \text{margin of error} = \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Maximum marks: 2

$$\alpha = 0.05 \quad \beta = 0.2$$

$$n = 10$$

$$\bar{X} = 8.5$$

$$\Rightarrow \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$$

$$\Rightarrow \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = \pm 1.96 \left(\frac{0.2}{\sqrt{10}} \right) = \pm 0.12$$

29 **New Question: 2 points**

The probability of at least one head in three flips of a fair coin is

(Do NOT round your answer) $p = 0.5$

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \binom{3}{0} 0.5^0 0.5^3$$

Maximum marks: 2

30 **New Question: 2 points**

You are working on a data set which has 91 observations. Your research question is whether the population variance differs from 9. You calculate the sample standard deviation, which turns out

to be 3. The value of your test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{90(3^2)}{9}$$

Maximum marks: 2

$$\begin{array}{l} H_0: \sigma^2 = 9 \quad n = 91 \\ H_a: \sigma^2 \neq 9 \quad s = 3 \end{array} \left. \begin{array}{l} \rightarrow \chi^2 \text{ test} \\ df = 90 \end{array} \right\}$$

= 90

31 **New Question: 3 points**

😊

Our class requires students to submit at least 7 assignments out of the available 9 assignments, while assignments 6, 7, 8, and 9 are compulsory. At the beginning of the semester, you intended to submit exactly 7 assignments including the compulsory assignments. How many ways of choosing the assignments to submit did you have?

Fill in your

here.

→ Your choices here are assignments 1, 2, 3, 4, 5 (since 6-9 are compulsory)
Must choose 3 to get 7 assignments

$$\rightarrow \binom{5}{3}$$

Maximum marks: 3

32 **New Question: 2 points**

A teacher claims that the students performed very similarly in a recent test and that the standard deviation of all students' scores is less than or equal to 0.5 (the test score ranges from 0 to 100).

You want to test this claim, your null hypothesis should be $H_0: \sigma^2 \leq$

$$0.5^2 = 0.25$$

Maximum marks: 2

33 New Question: 3 points

Consider that the distribution of your arrival time to your university classes follows a normal distribution with a mean of -5 minutes (i.e, you often come to your classes 5 minutes before the lecture time) and a standard deviation of 0.4 minutes.

Knowing that your next statistics class starts at 10:15 a.m. and given a 97.5% chance, what is the ~~start~~ time you will ~~leave~~ arrive to your statistics class?

$$\mu = -5$$

$$\sigma = 0.4$$

→ X follows normal dist.

$$P = 0.975$$

Fill in your answer here. (Write the answer in the format of xx:yy.)

$$\rightarrow P(X = x) = 0.975 \rightarrow P\left(z = \frac{x - \mu}{\sigma}\right)$$

Maximum marks: 3

34 New Question: 2 points

Consider that the distribution of your arrival time to your university classes follows a normal distribution with a mean of -5 minutes (i.e, you often come to your classes 5 minutes before the lecture time) and a standard deviation of 0.4 minutes.

Knowing that your next statistics class starts at 10:15 a.m., what is the probability that you will arrive later than 10:15 a.m.?

$$= P\left(z = \frac{x - (-5)}{0.4}\right) = P\left(z = \frac{x + 5}{0.4}\right) = 0.975$$

$$\Rightarrow \frac{x + 5}{0.4} = 1.96$$

$$\Rightarrow x = -4.216$$

→ 4 min early
⇒ 10:11 am

Fill in your answer here.

$$P(X > 0) = P\left(z > \frac{5}{0.4}\right) = P(z > 12.5)$$

$$= 1 - P(z \leq 12.5) \approx 1 = 0$$

Maximum marks: 2

35 ~~New Question: 2 points~~

⇒ Threshold values for grades are lowered by 2 points because of this question!

An animal shelter has 52 dogs and cats. 45% of the shelter animals are dogs. 62% of the animals are female. 45% of the animals are dogs or female. → DOESN'T MAKE SENSE!

If choosing an animal from the shelter at random, find the probability that the animal is a female dog?

Fill in your answer here. (Write the answer in TWO decimal places

and separate the ones and tenths by ".", not ",",. Eg: 2.32, 3.54, etc.)

Maximum marks: 2

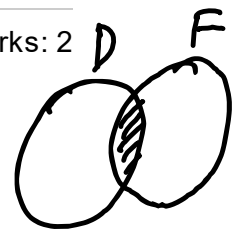
$$P(D) = 0.45 \rightarrow P(C) = 0.55 ; P(D \cup F) = 0.45$$

$$P(F) = 0.62 \rightarrow P(M) = 0.38$$

$$P(D \cap F) = ? \rightarrow P(D \cap F) = P(D) + P(F) - P(D \cup F)$$

$$= 0.45 + 0.62 - 0.45$$

$$= 0.62$$



- 36 **New Question: 2 points** *Not present $\rightarrow NP \rightarrow P(NP) = 0.08$
Fail $\rightarrow F$ $P(F|NP) = 0.5$*

8% of students do not show up in their class at all in a semester and account for 50% of the exam failure. What's the probability that a student does not at all show up in class and ends up failing that class that semester?

$$P(NP \cap F) = P(F|NP) \cdot P(NP) = 0.5 (0.08)$$

Fill in your answer here. (Write the answer in TWO decimal places and separate the ones and tenths by ".", not ",",. Eg: 2.32, 3.54, etc.)

Maximum marks: 2

- 37 **New Question: 2 points** *$n=16; \bar{X}=3, s=0.8$*

A random sample with a sample size of 16 has a mean of 3 and standard deviation of 0.8. What is the 95% confidence interval for the population mean?

$$\alpha=0.05$$

$$\bar{X} \pm t_{0.05} \left(\frac{0.8}{\sqrt{16}} \right) = 3 \pm 2.131 (0.2) = 3 \pm 0.4262$$

The interval is here. (Write the answer in a parenthesis (.), each number with TWO decimal places Eg: (2.32, 3.54), etc.)

Maximum marks: 2

- 38 **New Question: 2 points**

A sample of 35 dogs (Shiba breed) gives a mean weight of 10.5kg and a coefficient of variation of 18.1%. What is the variance of this sample?

$$\frac{\text{std}}{\text{mean}} \cdot 100\% = 18.1\% \Rightarrow \text{std} = 0.181 (10.5) = 1.9005 \rightarrow \text{variance} = (1.9005)^2 = 3.61$$

Fill in your answer here. (Write the answer in TWO decimal places and separate the ones and tenths by ".", not ",",. Eg: 2.32, 3.54, etc.)

Maximum marks: 2

- 39 **New Question: 2 points**

The chance for winning a lottery follows a uniform distribution among all the tickets, which are numbered from 100 to 250. What is the chance of a ticket numbered 123 winning the lottery?

Fill in your answer %. (Write the answer in THREE decimal places and separate the ones and tenths by ".", not ",",. Eg: 2.321, 3.540, etc.)

$$\rightarrow n = (250 - 100 + 1) = 151 \Rightarrow 151 \text{ tickets}$$

Maximum marks: 2

$$\rightarrow \frac{1}{151} = 0.662\%$$

40 New Question: 2 points

$P(F) = 0.3$; $P(DP|F) = 0.15$; $1 - P(DP) = 0.55$
 $\rightarrow P(NF) = 0.7$ $\rightarrow P(DP) = 0.45$

It shows in a survey that the proportion of people who floss their teeth daily accounts for 30%. Among these people, 15% of these people have a dental problem. 55% of the people in the survey do not have a dental problem. What is the probability that a person who does not floss daily and has a dental problem?

$P(NF \cap DP) = P(DP) - P(F \cap DP) = 0.45 - P(DP|F)P(F)$
 $= 0.45 - 0.15(0.3) = 0.405$

Fill in your answer here. (Write the answer in THREE decimal

places and separate the ones and tenths by “.”, not “,”. Eg: 2.321, 3.540, etc.)

$P(\text{Dental Problem}) = P(\text{Not Floss} \cap \text{Dental Problem}) + P(\text{Floss} \cap \text{Dental Problem})$
 Maximum marks: 2

41 New Question: 4 points

You play a game that can simulate different scenarios for Lord of the Rings, which will result in either Merry or Pippin reaching Mount Doom to destroy the Ring. You play the Pippin scenario 31 times, which shows a sample average of 21 months and 18 days and a variance of 28 days to reach Mount Doom. For the 26 times you play the Merry scenario, the average time to reach Mount Doom is 15 months and 19 days and the variance is 14 days. (There is no guarantee in any scenario that the Ring actually gets destroyed...)

Two-tailed : $H_0: \sigma_1^2 = \sigma_2^2$

Is there a significant difference between the time variances to reach Mount Doom by Merry and Pippin, given a level of significance = 0.1?

Fill in your here. (Answer “Yes” or “No”)

Maximum marks: 4

42 New Question: 3 points

Over time, the average proportion of undergraduate students is 80%, and 20% are postgraduate students. The university wants to attract more postgraduate students, so they try a new marketing strategy for their postgraduate programs. A year after the new strategy, 60 out of the total 250 students in a faculty take a postgraduate program (which previously had the same 80:20 ratio between undergraduate and postgraduate students). The dean of the faculty would like to check if the new strategy worked for the faculty.

$H_0: p \leq 0.2$
 $H_a: p > 0.2$

$\bar{p} = \frac{60}{250} = 0.24$

Should you suggest to the dean that the program worked for the faculty with a significance level of 0.05?

Test stat: $z = \frac{0.24 - 0.2}{\sqrt{\frac{0.2(0.8)}{250}}} = 1.58 \rightarrow p\text{-value} = 0.0571 > \alpha = 0.05$
 \rightarrow cannot reject H_0

Fill in your answer here. (Answer “Yes” or “No”)

Pippin has variance > Merry \rightarrow Pippin

Maximum marks: 3

is population 1

$\Rightarrow n_1 = 31, s_1^2 = 28$
 $n_2 = 26, s_2^2 = 14$
 $\alpha = 0.1$

$F = \frac{s_1^2}{s_2^2} = \frac{28}{14} = 2$
 num. df = $n_1 - 1 = 30$
 den. df = $n_2 - 1 = 25$

$\Rightarrow F_{\frac{\alpha}{2}, \text{num. df}, \text{den. df}} \Rightarrow F > \text{critical value}$
 $F_{0.05} = 1.92 \rightarrow \text{Reject } H_0$
 (From earlier) 19/24

43 New Question: 2 points

For a lower-tail test with a test statistic value of $z = -2.05$, what is the p-value for this test statistic?

Fill in your answer here. (Do NOT round your answer, type in what you have found from the appropriate table provided in the Support Material)

Maximum marks: 2

44 New Question: 2 points

20% of households start to have Christmas decoration before December. A sample of 15 households is selected. What is the chance that at least two of these 15 households have Christmas before December?

Fill your answer % here. (Write the answer in TWO decimal places and separate the ones and tenths by ".", not ",",. Eg: 2.32, 3.54, etc.)

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X=0) + P(X=1)]$$

$$= 1 - \left[\binom{15}{0} 0.2^0 0.8^{15} + \binom{15}{1} 0.2^1 0.8^{14} \right] = 0.8329$$

45 New Question: 3 points

32 batches of crystal cooked by Mr. White (refer to this as sample 1) shows a 97.3% average purity. Historical records show that out of all batches Mr. White has ever cooked, it is known that his purity standard deviation is 0.5%.

18 batches of crystal cooked by Mr. Pinkman (refer to this as sample 2) shows a 95% average purity. Historical records show that out of all batches Mr. Pinkman has ever cooked, it is known that his purity standard deviation is 2%.

A seller claims that with a 95% confidence level that Mr. Pinkman's crystal is as pure as Mr. White's. Is this claim true?

Fill in your answer here. (Answer "Yes" or "No")

$$\begin{aligned} \rightarrow H_0: \mu_1 - \mu_2 &= 0 \rightarrow D_0 = 0 \\ H_a: \mu_1 - \mu_2 &\neq 0 \end{aligned}$$

$$n_1 = 32; \bar{X}_1 = 0.973; \sigma_1 = 0.005$$

$$n_2 = 18; \bar{X}_2 = 0.95; \sigma_2 = 0.02$$

$$\text{Test stat: } z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.023}{\sqrt{\frac{0.05^2}{32} + \frac{0.02^2}{18}}} = 4.795 \rightarrow p\text{-value} \rightarrow 0$$

$\rightarrow \text{Reject } H_0$

or critical val $z_{0.975} = 1.96 \rightarrow z > \text{critical value}$
 $\rightarrow \text{Reject } H_0$

46 **New Question: 2 points**

You are taking part in a live TV program with a friend. Each of you is in a separate room and has to answer a multiple-choice question with four possible choices (only one choice is the correct answer). If both of you answer correctly, you win the final prize.

However, neither of you knows the answer and you both have to randomly guess it. What is the chance that you two will win the final prize?

Fill in your answer % here. (Write the answer in TWO decimal places and separate the ones and tenths by ".", not ",", Eg: 2.32, 3.54, etc.)

$$P(\text{You answer correctly}) = 0.25$$

$$P(\text{Your friend answers correctly}) = 0.25$$

→ independent events
Maximum marks: 2

$$\rightarrow P(\text{You} \cap \text{Your friend}) = 0.25(0.25)$$

47 **New Question: 3 points**

A university survey shows that 35% of the students live in the dorms, the rest live in private apartments. 30% of the students living in private apartments come from the same city as the location of the university. Only 1% of the students living in the dorms come from the same city.

If a student comes from that city, what is the chance that that student lives in one of the dorms?

Fill in your answer % here. (Write the answer in TWO decimal places and separate the ones and tenths by ".", not ",", Eg: 2.32, 3.54, etc.)

$$P(D) = 0.35; P(C|A) = 0.3; P(C|D) = 0.01$$

$$\rightarrow P(A) = 0.65$$

$$P(D|C) = \frac{P(D \cap C)}{P(C)} = \frac{P(C|D) \cdot P(D)}{P(C|D) \cdot P(D) + P(C|A) \cdot P(A)} = \frac{0.01(0.35)}{0.01(0.35) + 0.3(0.65)}$$

Maximum marks: 3

$$= 0.0176$$

48 **New Question: 4 points**

You get to travel to a fantasy world and bring back anything you like. You decide to enter the world created by George R. R. Martin and visit the Targaryens, who claimed that their blood was 100% pure. You sneak around (safely) and manage to get 16 blood samples from 16 different Targaryens. You are now back to the real world and ready to test the Targaryens' claim. From this hair sample, you get an average blood purity of 94% with a standard deviation of 8%. With a confidence level of 99%, would you reject the claim?

Fill in your answer here. (Answer "Yes" or "No")

From question 26 above $p\text{-value} < 0.01$

Maximum marks: 4

$$\alpha = 0.01 \rightarrow \text{Reject } H_0$$

49 **New Question: 2 points**

The weight of male football players is believed to follow a normal distribution. A random sample from a number of players gives a standard deviation of 5kg. Based on this sample, a 95% confidence interval estimate of the population variance is determined to be (10.93, 103.55). How many players were included in this sample, 8 or 16?

Fill in your here.

$s^2 = 25$;

Try $n = 8 \rightarrow \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{\alpha/2}} \rightarrow \frac{7(25)}{16.013} \leq \sigma^2 \leq \frac{7(25)}{1.69}$
 Maximum marks: 2
 \downarrow 10.93 \downarrow 103.55
 \rightarrow Correct!

50 **New Question: 3 points**

You are working with two different samples of equal size from the same population and get two different sample means (the second sample has a bigger sample mean). The first sample gives a larger sample standard deviation. Which sample would give a larger deviation from its sample mean when calculating a 95% confidence interval estimate for the population mean?

\Rightarrow margin of error

Fill in your answer here: sample. (Answer "first" or "second")

Maximum marks: 3

51 **New Question: 2 points**

You are taking a random sample of size ⁴ with a known population standard deviation of 18. Find the chance that the sample mean will be within 12 of the population mean?

$\sigma = 18$

Fill in your answer here %. (Write the answer in TWO decimal places and separate the ones and tenths by ".", not ",",. Eg: 2.32, 3.54, etc.)

$\rightarrow P(-12 \leq \bar{X} - \mu \leq 12) = P\left(\frac{-12}{\sigma_{\bar{X}}} \leq \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \leq \frac{12}{\sigma_{\bar{X}}}\right)$ Maximum marks: 2
 $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{18}{\sqrt{4}} = 9$
 $\mu_{\bar{X}} = E(\bar{X}) = \mu$
 $\rightarrow \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = z$
 $= P(-1.33 \leq z \leq 1.33)$
 $= 0.9082 - 0.0918 = 0.8164$

50) $\bar{X} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$
 margin of error
 sample standard deviation $s >$ \rightarrow margin of error $>$
 deviation from sample mean
 since $s_1 > s_2 \rightarrow t_{\alpha/2} \cdot \frac{s_1}{\sqrt{n}} > t_{\alpha/2} \cdot \frac{s_2}{\sqrt{n}}$
 $t_{\alpha/2} \rightarrow$ same (since same $df = n-1$)
 $\sqrt{n} \rightarrow$ same

52 **New Question: 3 points**

You are comparing the proportions of students getting a job within three months after their graduation from two different universities, and this is the information that you have gathered:

University 1	University 2
$n_1 = 160$	$n_2 = 100$
Students getting a job within 3 months = 40 $\rightarrow \bar{p}_1 = 0.25$	Students getting a job within 3 months = 25 $\rightarrow \bar{p}_2 = 0.25$

$$H_0: \bar{p}_1 - \bar{p}_2 = 0$$

$$H_a: \bar{p}_1 - \bar{p}_2 \neq 0$$

Using a 0.05 level of significance, do you conclude that the proportions differ between the two universities?

Fill in your answer here. (Answer "Yes" or "No")

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2} = \frac{40 + 25}{160 + 100} = 0.25$$

Test stat: $z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 0$

Maximum marks: 3 $\rightarrow p$ -value (2-tailed) = 1

53 **New Question: 3 points**

You are comparing the proportions of students getting a job within three months after their graduation from two different universities, and this is the information that you have gathered:

University 1	University 2
$n_1 = 160$	$n_2 = 100$
Students getting a job within 3 months = 40	Students getting a job within 3 months = 25

Using a 95% confidence interval, what is the lower limit of the estimated interval for the difference in the proportion of the two universities?

Fill in your answer here. (Write the answer in THREE decimal places and separate the ones and tenths by ".", not ",", Eg: 2.321, 3.540, etc.)

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}} = -1.96 \sqrt{\frac{0.25(0.75)}{160} + \frac{0.25(0.75)}{100}}$$

Maximum marks: 3

54 **New Question: 2 points**

The recorded price of a take-away meal in a city is:

Price (\$)	0 - 4	5 - 7	8 - 10	11+
Frequency	15	32	28	25

What is the median price in this sample?

Fill in your here.

$$x_1 \dots x_{15} \quad x_{16} \dots x_{47} \quad x_{48} \dots x_{75}$$

$$n = 15 + 32 + 28 + 25 = 100$$

n is even \rightarrow median = avg. of x_{50} and x_{51}
 \Rightarrow 8-10 group

Maximum marks: 2

55 **New Question: 2 points**

You are tossing a fair coin 7 times with the probability of getting heads is 0.5. What is the standard deviation of this distribution?

→ binomial dist.

Fill in your answer

here. (Write the answer in TWO decimal places

and separate the ones and tenths by ".", not ",". Eg: 2.32, 3.54, etc.)

Maximum marks: 2

$$\begin{aligned}\text{variance} &= np(1-p) \\ &= 7(0.5)0.5 = 1.75 \\ \rightarrow \text{standard deviation} &= 1.32\end{aligned}$$