MAT 2377C

Final practice

20 December 2018 **Professor: Mohsen Rezapour**
Time: 180 minutes

Student Number: ________________

Family Name: ________________

First Name: ________________

- This is a closed book exam.
- Your package includes the title page, five pages with questions, a formula sheet and tables.
- Only the calculators TI 30, TI 34, Casio fx-260 and Casio fx-300 are allowed.
- At the end of exam you need to submit the complete exam booklet.
- **Record your answer to each question in the table below.**
- Number of questions: **22**.
- If you see any error in this exam, please report it on your paper.

```
Cellular phones, unauthorized electronic devices or course notes
(unless an open-book exam) are not allowed during this exam.
Phones and devices must be turned off and put away in your bag.
Do not keep them in your possession, such as in your pockets.
If caught with such a device or document, the following may occur:
you will be asked to leave immediately the exam and academic fraud
allegations will be filed which may result in you obtaining a
0 (zero) for the exam.

By signing below, you acknowledge that you have ensured that you
are complying with the above statement.

X___________________________________
```

Page for reporting errors

If you see any error in this exam, please report it in this page.

- **Record your answer to each question in the table below.**
- Shade **ONE** letter for each question. A question with more than one shading answers will not be marked.

Question	a	b	c	d	e
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					

GOOD LUCK !!!

Q1. Suppose that for a very large shipment of integrated-circuit chips, the probability of failure for any one chip is 0.09. Find the probability that at most 2 chips fail in a random sample of size 20. (The numbers are rounded down to the second decimal place).

(a) 0.13 (b) 0.90 (c) 0.73 (d) 0.20 (e) none of the preceding

Solution to Q1:

X has binomial distribution with parameters $n = 20$ and $p = 0.09$. So

$$P(X \leq 2) \;=\; \binom{20}{0}(.09)^0(.91)^{20} + \binom{20}{1}(.09)^1(.91)^{19} + \binom{20}{2}(.09)^2(.91)^{18} = 0.73$$

Q2. Assume that in a box containing 100 items, five of them are defective. We select 20 items without replacement. What is the probability that at most one of them will be defective?

(a) 0.319 (b) 0.420 (c) 0.453

(d) 0.739 (e) none of the preceding

Solution to Q2:

Let X be the number of defective items. We have

$$P(X=0) + P(X=1) = \frac{\binom{5}{0}\binom{95}{20}}{\binom{100}{20}} + \frac{\binom{5}{1}\binom{95}{19}}{\binom{100}{20}} = 0.739.$$

Q3. A company which produces a particular drug has three factories, A, B and C. 30% of the drug are made in factory A, 50% in factory B, and 20% in factory C. Suppose that 95% and 97% of the drugs produced by the factory A and B, respectively, meet specifications while only 75% of the drugs produced by the factory C meet specifications. If I buy the drug and it meet specifications, what is the probability that it produced by factory B?

(a) 0.52 (b) 0.95 (c) 0.75

(d) 0.7 (e) none of the preceding

Solution to Q3:

M - meets specifications, A - produced by A, B - produced by B, and C - produced by C. Given: $P(A) = 0.3$, $P(B) = 0.5$ $P(C) = 0.2$, $P(M|A) = 0.95$, $P(M|B) = 0.97$, and $P(M|C) = 0.75$. To find: $P(M)$. Use the total probability rule:

$$P(M) = P(M|A)P(A) + P(M|B)P(B) + P(M|C)P(C) = 0.3 * 0.95 + 0.5 * 0.97 + 0.2 * 0.75 = 0.92$$

$$P(B|M) = \frac{P(M|B)P(B)}{P(M)} = \frac{0.5 * 0.97}{0.92} = 0.5271739$$ The answer a) was corrected

Q4. Assume that X is a discrete random variable with the following probability function:

$$f_X(x) = \begin{cases} k & \text{if } x = 0 \text{ or } 1 \\ \frac{14}{25} & \text{if } x = 3 \\ 0 & \text{otherwise.} \end{cases}$$

What is $F(2.7)$, where F is cumulative distribution function.

(a) 11/25 (b) 11/50 (c) 22/25

(d) 22/50 (e) None of the preceding.

Solution to Q4:

First we need to obtain k. $P(X = 0) + P(X = 1) + P(X = 3) = 1$, so that $k + k + 14/25 = 1$. Therefore, $k = 11/50$. Since $F(2.7) = P(X \leq 2.7) = P(X = 0) + P(X = 1) = 11/50 + 11/50 = 22/50$, the answer is a)

Q5. It is known that a particular company produces 30% defective items. 10 items are selected randomly. Calculate probability that at most two of them are defective.

(a)	0.3828	(b)	0.0014	(c)	0.2451
(d)	0.7549	(e)	None of the preceding.		

Solution to Q5:

Let X has binomial distribution with $n = 10$ and $p = .3$. Compute $P(X \leq 2)$.

$$P(X = 0) + P(X = 1) + P(X = 2)$$

Use the binomial formula.

Q6. Assume that times of failures of a computer network can be modeled using a Poisson process with mean 10 failures per year. What is the standard deviation of the waiting time to the first failure? (Time is measured as the part of a year, for example 0.5 means half a year).

(a)	0.222	(b)	0.1	(c)	0.333
(d)	0.471	(e)	none of the preceding		

Solution to Q6:

Waiting time to the first failure, X, is exponential $\lambda = 10$. Standard deviation is $\sqrt{\text{Var}(X)} = \sqrt{1/\lambda^2} = \sqrt{1}/10 = 0.1$.

Q7. Assume that X is a continuous random variable with the density f_X:

$$f_X(x) = \begin{cases} Ke^{-3x} & \text{if } x \geq -\ln(3)/3, \\ 0 & \text{otherwise.} \end{cases}$$

The probability $P(0 < X < 1)$ is

(a)	$\frac{\ln 3}{3}$	(b)	$-\frac{\ln 3}{3}$	(c)	$\frac{1}{3}(1 - e^{-3})$
(d)	$1 - e^{-3}$	(e)	None of the preceding.		

Solution to Q7:

First we should obtain K. $1 = \int_a^\infty f_X(x)\, dx = \int_{-\ln(3)/3}^\infty K \exp(-3x)\, dx = -\frac{1}{3} \exp(-3x)|_{x=-\ln(3)/3}^\infty \times K = K$ so that $K = 1$. Therefore, $P(0 < X < 1) = \int_0^1 f(t)dt = \int_0^1 e^{-3x}dx = \frac{-1}{3}(e^{-3} - 1)$

Q8. To estimate the mean of a population, the random sample X_1, X_2, X_3, X_4, X_5 are taken from the population with mean μ and variance σ^2 and the following estimators for mean are proposed: $\hat{\mu}_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$, $\hat{\mu}_2 = \frac{X_1 + X_2 + X_3}{3}$, $\hat{\mu}_3 = \frac{X_1 + X_2 + 3X_5}{5}$, $\hat{\mu}_4 = \frac{X_1 + X_2 + 2X_3}{5}$. Which statement is true?

(a) $\hat{\mu}_1$ is not an unbiased estimator.

(b) $\hat{\mu}_2$ is the most efficient estimator.

(c) $\hat{\mu}_3$ is more efficient that $\hat{\mu}_2$.

(d) $\hat{\mu}_4$ is an unbiased estimator.

(e) None of the preceding.

Solution to Q8:

We have

(a) Since $E(\hat{\mu}_1) = \frac{1}{5}(E(X_1)+\ldots+E(X_5)) = \mu$, $E(\hat{\mu}_2) = \frac{1}{3}(E(X_1)+\ldots+E(X_3)) = \mu$, $E(\hat{\mu}_3) = \frac{1}{5}(E(X_1)+E(X_2)+3E(X_5)) = \mu$, $E(\hat{\mu}_4) = \frac{1}{5}(E(X_1)+E(X_2)+2E(X_3)) = \frac{4}{4}\mu$, $\hat{\mu}_1, \hat{\mu}_2$, and $\hat{\mu}_3$ are unbiased estimator and $\hat{\mu}_4$ is biased estimator.

(b) $\sigma^2_{\hat{\mu}_1} = \frac{1}{5^2}(\sigma^2_{X_1}+\ldots+\sigma^2_{X_5}) = \sigma^2/5$, $\sigma^2_{\hat{\mu}_2} = \frac{1}{5^2}(\sigma^2_{X_1}+\sigma^2_{X_1}+\sigma^2_{X_3}) = \sigma^2/3$, $\sigma^2_{\hat{\mu}_3} = \frac{1}{5^2}(\sigma^2_{X_1}+\sigma^2_{X_1}+3^2\sigma^2_{X_5}) = \frac{11}{3}\sigma^2$, $\sigma^2_{\hat{\mu}_4} = \frac{1}{5^2}(\sigma^2_{X_1}+\sigma^2_{X_1}+2^2\sigma^2_{X_3}) = \frac{6}{3}\sigma^2$,
Since $\hat{\mu}_1$ has the smallest variance it is the most efficient estimator. Therefore, e is correct answer.

Marking scheme for Q8:

1 point for each correct answer of the first part, 1 point for each correct answer of the second part, 1 point for the correct answer to the third part. Total - 9 points.

Q9. The air pressure in a randomly selected tire put on a certain model new car is normally distributed with mean value 31 lb/in^2 and standard deviation 0.5 lb/in^2. What is the probability that the pressure for a randomly selected tire is between 30.5 and 31.5 lb/in^2.

(a) 0.6827 (b) 0.3173 (c) 0.5000

(d) 0.4245 (e) none of the preceding

Solution to Q9:

Let X be air pressure of a randomly selected tire. Then X has normal distribution with$\mu = 31$ and $\sigma = 0.5$. Therefore,

$$P(30.5 < X < 31.5) = P\left(\frac{30.5-31}{0.5} < Z < \frac{31.5-31}{0.5}\right) = P(-1 < Z < 1) = 0.6827$$

Q10. Assume that we have a sample $X_1, \ldots, X_{100}$ from an arbitrary distribution with mean $\mu = 1$ and variance $\sigma^2 = 4$. The approximate probability that the sample mean exceeds 1.5, is:

(a) 0.99379 (b) 0.00621 (c) 0.5000 (d) 0.7742 (e) none of the preceding

Solution to Q10:

$$P(\bar{X} > 1.5) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{1.5-1}{2/10}\right) = P(Z > 2.5) = 1 - 0.993790 = 0.00621$$

Q11. Assume that we have a sample of size 10 from a population $N(4,9)$. Denote by $\overline{X}$ and S^2, the sample mean and sample variance, respectively. Find c such that

$$P\left(\frac{\overline{X}-4}{S/\sqrt{10}} \geq c\right) = 0.01.$$

(a)	1.833	(b)	2.821	(c)	1.645
(d)	2.424	(e)	none of the preceding		

Solution to Q11:

Equivalent statement: find c such that

$$P\left(\frac{\overline{X}-4}{S/\sqrt{10}} \geq c\right) = 0.01.$$

We have that $\dfrac{\overline{X}-4}{S/\sqrt{10}}$ has Student distribution with $n-1=9$ degrees of freedom. From the table we read that $P(T > 2.821) = 0.01$, thus $c = 2.821$.

Q12. Assume we have a sample of size $n=5$: $5, 34, 12, 10, 4$. The sample mean, the sample variance and the sample median are, respectively:

(a)	13, 149 and 10	(b)	13, 119.2 and 12	(c)	12, 119.2 and 5
(d)	13, 12.2 and 4.5	(e)	none of the preceding		

Solution to Q12:

$\overline{x} = \sum x_i/n = 65/5 = 13$ and

$$s^2 = \frac{(\sum x_i^2) - (\sum x_i)^2/n}{n-1} = \frac{1141-(65)^2/5}{4} = 149$$

Q13. The engineer measures $n=25$ pieces of steel and obtains $\bar{x}=6$. The weight follows normal distribution with known variance $\sigma^2 = 16$. The two-sided 99% confidence interval for the mean is:

(a)	(-0.272,12.272)	(b)	(3.250,8.750)
(c)	(4.432,7.568)	(d)	(3.940,8.060)
(e)	None of the preceding.		

Solution to Q13:

$$(\bar{x} - z_{\alpha/2}\sigma/\sqrt{n}, \bar{x} + z_{\alpha/2}\sigma/\sqrt{n}) = (6 - 2.575 * 4/5, 6 + 2.575 * 4/5) = (3.94, 8.06)$$

Q14. A machine is producing metal pieces that are cylindrical in shape. A sample of pieces is taken and the diameters are

```
1.01 0.97 1.03 1.04 0.99 0.98  0.99 1.04 1.03 1.01
```

Find the two-sided 99% confidence interval for the true mean diameter. Assume that the population is normally distributed.

(a) [0.989,1.022] (b) [0.983,1.035] (c) [0.991,1.034]

(d) [.987,1.024] (e) none of the preceding.

Solution to Q14:

The sample mean and standard deviation are $\overline{x} = 1.009$ and $s = 0.0256$, respectively.

We have $t_{.005} = 3.250$ with $10 - 1$ degrees of freedom. So

$$\overline{x} \pm 3.250\,\frac{s}{\sqrt{n}} = [0.983, 1.035].$$

Q15. The engineer measures $n = 25$ pieces of steel and obtains $\bar{x} = 6$. The weight follows normal distribution with known variance $\sigma^2 = 16$. He wants to test $H_0 : \mu = 5$ against $H_1 : \mu > 5$. The p-value for the test is:

(a) 0.0500 (b) 0.1057 (c) 0.8943

(d) 1.000 (e) None of the preceding.

Solution to Q15:

$$P(\bar{X} > 6) = P\left(Z > \frac{6-5}{4/5}\right) = P(Z > 1.25) = 1 - 0.8943 = 0.1057.$$

Q16. A company claims that the mean deflection of a piece of steel which is 10 feet long, is equal to 0.012. A buyer suspects that it is bigger than 0.012. The following data has been collected:

0.0132 0.0138 0.0108 0.0126 0.0136 0.0112 0.0124 0.0116 0.0127 0.0131

Assume normality. *Hint:* $\sum_{i=1}^{n} x_i^2 = 0.0016$. The p-value for the appropriate one-sided test and the decision are:

(a) $p \in (0.05, 0.1)$. Reject H_0 for $\alpha = 0.05$. (b) $p \in (0.05, 0.1)$. Do not reject H_0 for $\alpha = 0.05$.

(c) $p \in (0.1, 0.25)$. Reject H_0 for $\alpha = 0.05$. (d) $p \in (0.1, 0.25)$. Do not reject H_0 for $\alpha = 0.05$.

(e) None of the preceding.

Solution to Q16:

The estimated variance is

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2\right) = 1.02 * 10^{-6}.$$

The observed mean is $\bar{x} = 0.0125$. We calculate the p-value:

$$\begin{aligned} P\left(\bar{X} > 0.0125\right) &= P\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} > \frac{0.0125 - 0.012}{\sqrt{1.02 * 10^{-6}/10}}\right) \\ &= P(t_9 > 1.5638) \in (0.05, 0.1). \end{aligned}$$

Do not reject H_0 for $\alpha = 0.05$.

Q17. A company is currently using titanium alloy rods it purchases from supplier A. A new supplier (supplier B) approaches the company and offers the same quality (at least according to supplier Bs claim) rods at a lower price. The company is certainly interested in the offer. At the same time, the company wants to make sure that the safety of their product is not compromised. The company randomly selects ten rods from each of the lots shipped by suppliers A and B and measures the yield strengths of the selected rods. The observed sample mean and sample standard deviation are 651 MPa and 2 MPa for suppliers A rods, respectively, and the same parameters are 657 MPa and 3 MPa for supplier Bs rods. The company tests the hypotheses $H_0 : \mu_A = \mu_B$ against $\mu_A \neq \mu_B$. For $\alpha = 0.05$ the decision is:

(a) Reject H_0; (b) Do not reject H_0;

(c) none of the preceding

Assume equal variances.

Solution to Q17:

I added "assume equal variances" statement so that you can use case 2. I also modified the statement, to be in line with what I taught.

Q18. In an effort to compare the durability of two different types of sandpaper, 10 pieces of type A sandpaper were subjected to treatment by a machine which measures abrasive wear. Eleven pieces of type B sandpaper were subjected to the same treatment.

x_i	27	26	24	29	30	26	27	23	28	27	
y_i	24	23	22	27	24	21	24	25	24	23	20

Note: $\sum_i x_i = 267$, $\sum_i y_i = 257$, $\sum_i x_i^2 = 7169$, $\sum_i y_i^2 = 6041$. From this you can easily compute sample means and sample variances, for example

$$s_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}) = \frac{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}.$$

Assuming normality and equality of variance, we want to test for equality of mean abrasive wear. The appropriate p-value is:

(a) $p < 0.002$ (b) $p > 0.2$ (c) $p \in (0.05, 0.1)$

(d) $p \in (0.1, 0.2)$ (e) None of the preceding

Solution to Q18:

This is two sample test. $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$. We compute $S_1^2 = 4.45$, $S_2^2 = 3.65$, $S_p^2 = 4.03$, $\bar{X} = 26.7$l $\bar{Y} = 23.36$. The p-value is

$$2P\left(\bar{X} - \bar{Y} > 3.34\right) = 2P\left(\frac{\bar{X} - \bar{Y}}{S_p\sqrt{1/n + 1/m}} > \frac{3.34}{\sqrt{4.03}\sqrt{1/10 + 1/11}}\right) = 2P(t_{19} > 3.8037) < 0.002$$

Q19. The following output was produced with `t.test` command in `R`

```
One Sample t-test

data:  x
t = 32.9198, df = 999, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.9644462 1.0867155
sample estimates:
mean of x
 1.025581
```

Based on this output, which statement is correct:

(a) If the significant level of the test is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;

(b) If the significant level of the test is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu \neq 0$;

(c) If the significant level of the test is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;

(d) If the significant level of the test is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu < 0$;

(e) None of the preceding

Q20. Regression methods were used to analyze the data from a study investigating the relationship between roadway surface temperature in F (x) and pavement defection (y). Summary quantities were $n = 33$, $\sum y_i = 1124$, $\sum y_i^2 = 41998$, $\sum x_i = 1124$, $\sum x_i^2 = 41086$, $\sum x_i y_i = 41355$. The slope coefficient is:

(a) 0.9036 (b) 0.6854 (c) 0.6014 (d) 0.4045 (e) none of the preceding

Solution to Q20:

We have $S_{xy} = 3752.09$, $S_{xx} = 4152.018$, $S_{yy} = 3713.88$;

$$b_1 = \frac{S_{xy}}{S_{xx}}, \quad b_0 = \bar{y} - b_1 \bar{x},$$

where

$$\begin{aligned} S_{xy} &= \sum_{i=1}^{n} x_i y_i - \frac{1}{n}(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i) \\ S_{xx} &= \sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2 \\ S_{yy} &= \sum_{i=1}^{n} y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2 \end{aligned}$$

So $b_1 = 0.903643$

Q21. Consider the data from Question 20. The 95% prediction interval for the mean response $\mu_{Y|x_0}$ when $x_0 = 20$ is:

(a) (16.254,25.265)

(b) (17.265,38.245)

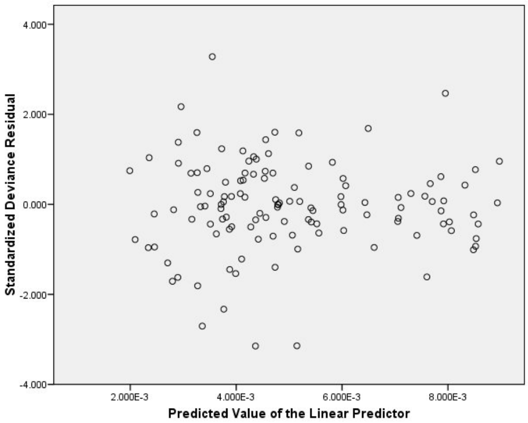

FIGURE 1. Plot of residuals against fitted value

(c) (20.107,23.697)

(d) (19.256,31.325)

(e) None of the preceding

Solution to Q21:

$S^2 = \frac{SSE}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2} = 3.2295\,,\ t_{\alpha/2} = 2.045$ The prediction interval is $\hat{y_0} \pm t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$
Answer c).

Q22. Figure 1 is the plot of the residuals against the fitted values in a regression model. Which one of the following statements about this figure is true.

(a) This plot confirms that the residuals do not have zero mean.

(b) This plot shows that the residuals have constant variance.

(c) This plot confirms that the residuals have zero mean.

(d) This plot confirms that the residuals have normal distribution.

(e) None of the preceding

Solution to Q22:

Answer c).

This is the last question

MAT 2377
Final Exam Formula Sheet

- Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Conditional probability of A given B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Total probability rule:

$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n) \end{aligned}$$

- Bayes' rule

$$P(A_r|B) = \frac{P(B|A_r)P(A_r)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

- Events A and B are independent if $P(A \cap B) = P(A)P(B)$
- Expected value of a discrete random variable X:

$$\mu = E(X) = \sum_x x f(x), \quad \text{where} \quad f(x) = P(X = x)$$

- Expected value of a function of a continuous random variable X:

$$E[h(X)] = \int h(x) f(x) dx, \quad \text{where } f(x) \quad \text{is a density}$$

- If a density f or a distribution F are given, then

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(u) du.$$

- Binomial random variable: X has Binomial distribution with parameters n and p:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

  Mean: np; Variance $np(1-p)$.
- Geometric random variable: X has Geometric distribution with parameter p:

$$P(X = k) = (1-p)^{k-1} p.$$

  Mean: $1/p$.
- Poisson random variable: X has Poisson distribution with parameter λt:

$$P(X = k) = e^{-(\lambda t)} \frac{(\lambda t)^k}{k!}.$$

  Mean: λ; Variance λ.
- Exponential random variable: X has exponential distribution with parameter $\beta = \frac{1}{\lambda}$:

$$f(x) = \frac{1}{\beta} e^{-x/\beta}, \quad x > 0.$$

  Mean: β.
- Gamma random variable: X has gamma distribution with parameters α and $\beta = \frac{1}{\lambda}$:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad x > 0.$$

  Mean: $\beta\alpha$.

  $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$, for $\alpha > 0$. $\Gamma(1) = 1$, and $\Gamma(n) = (n-1)!$ for positive integer n.

- Standardization: If X is a normal random variable with mean μ and variance σ^2, then

$$Z = \frac{X-\mu}{\sigma} \quad \text{has a standard normal distribution}$$

- Sample mean of the observations $x_1, \ldots, x_n$:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

- Sample variance of the observations $x_1, \ldots, x_n$:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right].$$

- Statistic used for confidence intervals and tests for a mean μ when σ is known:

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \text{ has a standard normal distribution}$$

- Statistic used for confidence intervals and tests for a mean μ when σ is unknown:

$$T_0 = \frac{\bar{X}-\mu}{S/\sqrt{n}} \text{ has a } T \text{ distribution with } n-1 \text{ d.f.}$$

- Statistic used for confidence intervals and tests for the means μ_1 and μ_2 of two independent populations with variances are known:

$$Z = \frac{\bar{X}-\bar{Y}}{\sqrt{(\sigma_1^2/n + \sigma_2^2/m)}} \text{ has a standard normal distribution}$$

- Statistic used for confidence intervals and tests for the means μ_1 and μ_2 of two independent normal populations with equal variances:

$$T_0 = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_p\sqrt{(1/n+1/m)}} \text{ has a } T \text{ distribution with } n+m-2 \text{ d.f.}$$

where

$$S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

- Statistic used for confidence intervals for the variance σ^2 of a normal population:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \text{ has a Chi-squared distribution with } n-1 \text{ d.f.}$$

Formulas for regression analysis:

- Parameter estimation:

$$b_1 = \frac{S_{xy}}{S_{xx}}, \quad b_0 = \bar{y} - b_1\bar{x},$$

$$\begin{aligned} S_{xy} &= \sum_{i=1}^{n} x_i y_i - \frac{1}{n}(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i) \\ S_{xx} &= \sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2 \\ S_{yy} &= \sum_{i=1}^{n} y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2 \end{aligned}$$

- Coefficient of determination $R^2 = 1 - SSE/SST$, where $SST = S_{yy}$ and $SSE = S_{yy} - b1S_{xy}$.

- Statistic used for confidence intervals and tests for the parameter β_0

$$T = \frac{b_0 - \beta_{00}}{S\sqrt{\sum_{i=1}^n x_i^2/(nS_{xx})}} \text{ has a } T \text{ distribution with } n-2 \text{ d.f.}$$

where

$$S^2 = \frac{SEE}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2}$$

- Statistic used for confidence intervals and tests for the parameter β_1

$$T = \frac{b_1 - \beta_{10}}{S\sqrt{1/(S_{xx})}} \text{ has a } T \text{ distribution with } n-2 \text{ d.f.}$$

where

$$S^2 = \frac{SEE}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2}$$

- Statistic used for confidence intervals for the mean response $\mu_{Y|x_0}$

$$T = \frac{\hat{y}_0 - \mu_{Y|x_0}}{S\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}}} \text{ has a } T \text{ distribution with } n-2 \text{ d.f.}$$

- Statistic used for confidence intervals for the single predicted value y_0

$$T = \frac{\hat{y}_0 - y_0}{S\sqrt{1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}}} \text{ has a } T \text{ distribution with } n-2 \text{ d.f.}$$

- Statistic used for ANOVA

$$f = SSR/S^2 \text{ has a } F\text{-distribution with } \nu_1 = 1 \text{ and } \nu_2 = n-2 \text{ d.f.}$$

where

$$SST = SSR + SSE, SST = S_{yy}\,.$$

Area under the normal curve ($P(Z < z)$, where Z has standard normal distribution).

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Area under the normal curve ($P(Z < z)$, where Z has standard normal distribution).

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Critical value of t-distribution $P(t > t_\alpha) = \alpha$.

v	0.40	0.30	0.20	0.15	0.10	0.05	0.025
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.538	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960

Critical value of Chi-squared distribution $P(X^2 > \chi^2_\alpha) = \alpha$.

ν	0.30	0.25	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.001
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412	6.635	7.879	10.827
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824	9.210	10.597	13.815
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.266
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668	13.277	14.860	18.466
5	6.064	6.626	7.289	9.236	11.070	12.832	13.388	15.086	16.750	20.515
6	7.231	7.841	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457
7	8.383	9.037	9.803	12.017	14.067	16.013	16.622	18.475	20.278	24.321
8	9.524	10.219	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.124
9	10.656	11.389	12.242	14.684	16.919	19.023	19.679	21.666	23.589	27.877
10	11.781	12.549	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588
11	12.899	13.701	14.631	17.275	19.675	21.920	22.618	24.725	26.757	31.264
12	14.011	14.845	15.812	18.549	21.026	23.337	24.054	26.217	28.300	32.909
13	15.119	15.984	16.985	19.812	22.362	24.736	25.471	27.688	29.819	34.527
14	16.222	17.117	18.151	21.064	23.685	26.119	26.873	29.141	31.319	36.124
15	17.322	18.245	19.311	22.307	24.996	27.488	28.259	30.578	32.801	37.698
16	18.418	19.369	20.465	23.542	26.296	28.845	29.633	32.000	34.267	39.252
17	19.511	20.489	21.615	24.769	27.587	30.191	30.995	33.409	35.718	40.791
18	20.601	21.605	22.760	25.989	28.869	31.526	32.346	34.805	37.156	42.312
19	21.689	22.718	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.819
20	22.775	23.828	25.038	28.412	31.410	34.170	35.020	37.566	39.997	45.314
21	23.858	24.935	26.171	29.615	32.671	35.479	36.343	38.932	41.401	46.796
22	24.939	26.039	27.301	30.813	33.924	36.781	37.659	40.289	42.796	48.268
23	26.018	27.141	28.429	32.007	35.172	38.076	38.968	41.638	44.181	49.728
24	27.096	28.241	29.553	33.196	36.415	39.364	40.270	42.980	45.558	51.179
25	28.172	29.339	30.675	34.382	37.652	40.646	41.566	44.314	46.928	52.619
26	29.246	30.435	31.795	35.563	38.885	41.923	42.856	45.642	48.290	54.051
27	30.319	31.528	32.912	36.741	40.113	43.195	44.140	46.963	49.645	55.475
28	31.391	32.620	34.027	37.916	41.337	44.461	45.419	48.278	50.994	56.892
29	32.461	33.711	35.139	39.087	42.557	45.722	46.693	49.588	52.335	58.301
30	33.530	34.800	36.250	40.256	43.773	46.979	47.962	50.892	53.672	59.702
40	44.165	45.616	47.269	51.805	55.758	59.342	60.436	63.691	66.766	73.403
50	54.723	56.334	58.164	63.167	67.505	71.420	72.613	76.154	79.490	86.660
60	65.226	66.981	68.972	74.397	79.082	83.298	84.58	88.379	91.952	99.608

Solutions to multiple choice questions:

Q1 $\longrightarrow$ c
Q2 $\longrightarrow$ d
Q3 $\longrightarrow$ a
Q4 $\longrightarrow$ a
Q5 $\longrightarrow$ a
Q6 $\longrightarrow$ b
Q7 $\longrightarrow$ c
Q9 $\longrightarrow$ a
Q10 $\longrightarrow$ b
Q11 $\longrightarrow$ b
Q12 $\longrightarrow$ a
Q13 $\longrightarrow$ d
Q14 $\longrightarrow$ b
Q15 $\longrightarrow$ b
Q16 $\longrightarrow$ b
Q17 $\longrightarrow$ a
Q18 $\longrightarrow$ a
Q19 $\longrightarrow$ b
Q20 $\longrightarrow$ a
Q21 $\longrightarrow$ c
Q22 $\longrightarrow$ c