MAT 2377
Probability and Statistics for Engineers

Practice Set

P. Boily (uOttawa)

Winter 2021

Based on course notes by Rafał Kulik

Q1. Two events each have probability 0.2 of occurring and are independent. The probability that neither occur is:

a) 0.64 b) 0.04 c) 0.2 d) 0.4 e) none of the preceding

Q2. Two events each have probability 0.2 and are mutually exclusive. The probability that neither occurs is:

a) 0.36 b) 0.04 c) 0.2 d) 0.6 e) none of the preceding

Q3. A smoke-detector system consists of two parts A and B. If smoke occurs then the item A detects it with probability 0.95, the item B detects it with probability 0.98 whereas both of them detect it with probability 0.94. What is the probability that the smoke will not be detected?

a) 0.01 b) 0.99 c) 0.04 d) 0.96 e) none of the preceding

Q4. Three football players will attempt to kick a field goal. Let A_1, A_2, A_3 denote the events that the field goal is made by player $1, 2, 3$, respectively. Assume that A_1, A_2, A_3 are independent and $P(A_1) = 0.5$, $P(A_2) = 0.7$, $P(A_3) = 0.6$. Compute the probability that exactly one player is successful.

a) 0.29 b) 0.21 c) 0.71 d) 0.79 e) none of the preceding

Q5. In a group of 16 candidates for laboratory research positions, 7 are chemists and 9 are physicists. In how many ways can one choose a group of 5 candidates with 2 chemists and 3 physicists?

Q6. There is a theorem of combinatorics that states that the number of permutations of n objects in which n_1 are alike of kind 1, n_2 are alike of kind 2, ..., and n_r are alike of kind r (that is, $n = n_1 + n_2 + \cdots + n_r$) is

$$\frac{n!}{n_1! \cdot n_2! \cdot \cdots \cdot n_r!}.$$

Find the number of different words that can be formed by rearranging the letters in the following words (include the given word in the count):

a)NORMAL b)HHTTTT c)ILLINI d)MISSISSIPPI

Q7. A class consists of 490 engineering students and 510 science students. The students are divided according to their marks:

	Passed	Failed
Eng.	430	60
Sci.	410	100

If one person is selected randomly, the probability that it failed if it was an engineering student?

a) 0.06 b) 0.12 c) 0.41 d) 0.81 e) none of the preceding

Q8. A company which produces a particular drug has two factories, A and B. 30% of the drug are made in factory A, 70% in factory B. Suppose that 95% of the drugs produced by factory A meet specifications while only 75% of the drugs produced by factory B meet specifications. If I buy a dose of the company's drug, what is the probability that it meets specifications?

a)0.81 b)0.95 c)0.75 d)0.7 e)none of the preceding

Q9. A medical research team wished to evaluate a proposed screening test for Alzheimer's disease. The test was given to a random sample of 450 patients with Alzheimer's disease; in 436 cases the test result was positive. The test was also given to a random sample of 500 patients without the disease; only in 5 cases was the result was positive. It is known that in Canada 11.3% of the population aged $65+$ have Alzheimer's disease. Find the probability that a person has the disease given that their test was positive (choose the closest answer).

a) 0.97 b) 0.93 c) 0.99 d) 0.07 e) none of the preceding

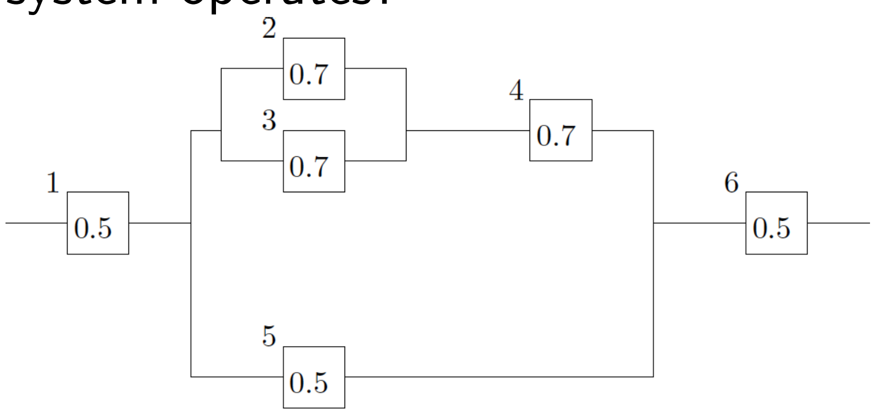
Q10. Twelve items are independently sampled from a production line. If the probability that any given item is defective is 0.1, the probability of at most two defectives in the sample is closest to ...

a)0.38748 b)0.9872 c)0.7361 d)0.8891 e)none of the preceding

Q11. A student can solve 6 problems from a list of 10. For an exam 8 questions are selected at random from the list. What is the probability that the student will solve exactly 5 problems?

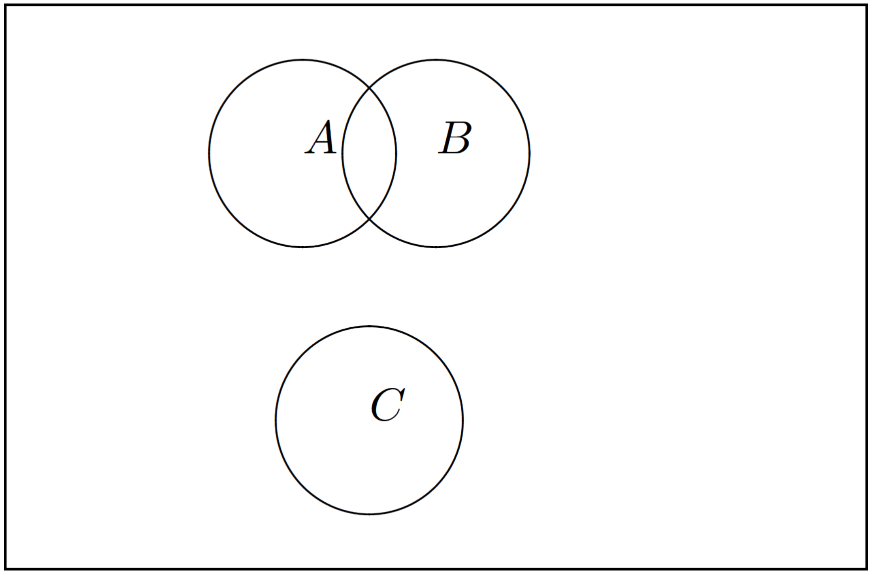
a) 0.98 b) 0.02 c) 0.28 d) 0.53 e) none of the preceding

Q12. Consider the following system with six components. We say that it is functional if there exists a path of functional components from left to right. The probability of each component functions is shown. Assume that the components function or fail independently. What is the probability that the system operates?



a) 0.1815 b) 0.8185 c) 0.6370 d) 0.2046 e) none of the preceding

Q13. Three events are shown in the Venn diagram below.



Shade the region corresponding to the following events:

a) A^c

b) $(A \cap B) \cup (A \cap B^c)$

c) $(A \cap B) \cup C$

d) $(B \cup C)^c$

e) $(A \cap B)^c \cup C$

Q14. Pieces of aluminum are classified according to the finishing of the surface and according to the finishing of edge. The results from 85 samples are summarized as follows:

	Edge	
Surface	excellent	good
excellent	60	5
good	16	4

Let A denote the event that a selected piece has "excellent" surface, and let B denote the event that a selected piece has "excellent" edge. If samples are elected randomly, determine the following probabilities:

a) $P(A)$ b) $P(B)$ c) $P(A^c)$

d) $P(A \cap B)$ e) $P(A \cup B)$ f) $P(A^c \cup B)$

Q15. If $P(A) = 0.1$, $P(B) = 0.3$, $P(C) = 0.3$, and events A, B, C are mutually exclusive, determine the following probabilities:

a)$P(A \cup B \cup C)$ b)$P(A \cap B \cap C)$ c)$P(A \cap B)$

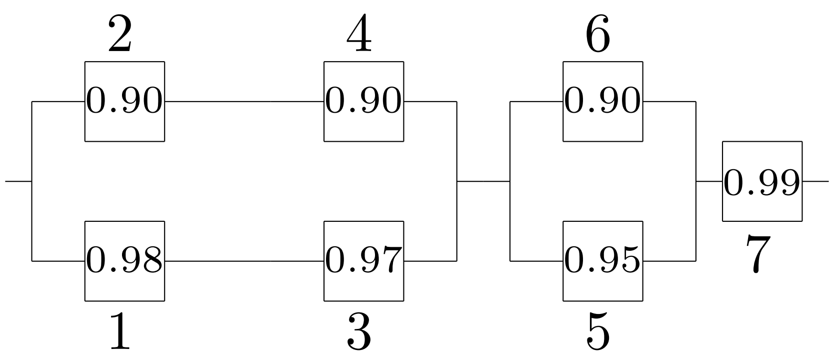
d)$P((A \cup B) \cap C)$ e)$P(A^c \cap B^c \cap C^c)$ f)$P[(A \cup B \cup C)^c]$

Q16. The probability that an electrical switch, which is kept in dryness, fails during the guarantee period, is 1%. If the switch is humid, the failure probability is 8%. Assume that 90% of switches are kept in dry conditions, whereas remaining 10% are kept in humid conditions.

a) What is the probability that the switch fails during the guarantee period?

b) If the switch failed during the guarantee period, what is the probability that it was kept in humid conditions?

Q17. The following system operates only if there is a path of functional device from left to the right. The probability that each device functions is as shown. What is the probability that the circuit operates? Assume independence.



Q18. An inspector working for a manufacturing company has a 95% chance of correctly identifying defective items and 2% chance of incorrectly classifying a good item as defective. The company has evidence that 1% of the items it produces are nonconforming (defective).

1. What is the probability that an item selected for inspection is classified as defective?
2. If an item selected at random is classified as non defective, what is the probability that it is indeed good?

Q19. Consider an ordinary 52-card North American playing deck (4 suits, 13 cards in each suit).

a) How many different $5-$card poker hands can be drawn from the deck?

b) How many different $13-$card bridge hands can be drawn from the deck?

c) What is the probability of an all-spade $5-$card poker hand?

d) What is the probability of a flush ($5-$cards from the same suit)?

e) What is the probability that a $5-$card poker hand contains exactly 3 Kings and 2 Queens?

f) What is the probability that a $5-$card poker hand contains exactly 2 Kings, 2 Queens, and 1 Jack?

Q20. Students on a boat send messages back to shore by arranging seven coloured flags on a vertical flagpole.

a) If they have 4 orange flags and 3 blue flags, how many messages can they send?

b) If they have 7 flags of different colours, how many messages can they send?

c) If they have 3 purple flags, 2 red flags, and 4 yellow flags, how many messages can they send?

Q21. The Stanley Cup Finals of hockey or the NBA Finals in basketball continue until either the representative team form the Western Conference or from the Eastern Conference wins 4 games. How many different orders are possible ($WWEEEE$ means that the Eastern team won in 6 games) if the series goes

a)4 games? b)5 games? c)6 games? d)7 games?

Q22. Consider an ordinary 52-card North American playing deck (4 suits, 13 cards in each suit), from which cards are drawn at random and without replacement, until 3 spades are drawn.

a) What is the probability that there are 2 spades in the first 5 draws?

b) What is the probability that a spade is drawn on the 6th draw given that there were 2 spades in the first 5 draws?

c) What is the probability that 6 cards need to be drawn in order to obtain 3 spades?

d) All the cards are placed back into the deck, and the deck is shuffled. 4 cards are then drawn from. What is the probability of having drawn a spade, a heart, a diamond, and a club, in that order?

Q23. A student has 5 blue marbles and 4 white marbles in his left pocket, and 4 blue marbles and 5 white marbles in his right pocket. If they transfer one marble at random from their left pocket to his right pocket, what is the probability of them then drawing a blue marble from their right pocket?

Q24. An insurance company sells a number of different policies; among these, 60% are for cars, 40% are for homes, and 20% are for both. Let A_1, A_2, A_3, A_4 represent people with only a car policy, only a home policy, both, or neither, respectively. Let B represent the event that a policyholder renews at least one of the car or home policies.

a) Compute $P(A_1)$, $P(A_2)$, $P(A_3)$, and $P(A_4)$.

b) From past data, we know that $P(B|A_1) = 0.6$, $P(B|A_2) = 0.7$, $P(B|A_3) = 0.8$. Given that a client selected at random has a car or a home policy, what is the probability that they will renew one of these policies?

Q25. An urn contains four balls numbered 1 through 4. The balls are selected one at a time, without replacement. A match occurs if ball m is the mth ball selected. Let the event A_i denote a match on the ith draw, $i = 1, 2, 3, 4$.

a) Compute $P(A_i)$, $i = 1, 2, 3, 4$.

b) Compute $P(A_i \cap A_j)$, $i, j = 1, 2, 3, 4$, $i \neq j$.

c) Compute $P(A_i \cap A_j \cap A_k)$, $i, j, k = 1, 2, 3, 4$, $i \neq j, i \neq k, j \neq k$.

d) What is the probability of at least 1 match?

Q26. The probability that a company's workforce has at least one accident in a given month is $(0.01)k$, where k is the number of days in the month. Assume that the number of accidents is independent from month to month. If the company's year starts on January 1, what is the probability that the first accident occurs in April?

Q27. A Pap smear is a screening procedure used to detect cervical cancer. Let T^- and T^+ represent the events that the test is negative and positive, respectively, and let C represent the event that the person tested has cancer.

The false negative rate for this test when the patient has the cancer is 16%; the false positive test for this test when the patient does not have cancer is 19%.

In North America, the rate of incidence for this cancer is roughly 8 out of 100,000 women. Based on these numbers, do you think that the Pap smear is an effective procedure? What factors influence your conclusion?

Q28. Of three different fair dice, one each is given to Elowyn, Llewellyn, and Gwynneth. They each roll the die they received.

Let $E = \{\text{Elowyn rolls a } 1 \text{ or a } 2\}$, $LL = \{\text{Llewellyn rolls a } 3 \text{ or a } 4\}$, and $G = \{\text{Gwynneth rolls a } 5 \text{ or a } 6\}$ be 3 events of interest.

a) What are the probabilities of each of E, LL, and G occurring?

b) What are the probabilities of any two of E, LL, and G occurring simultaneously?

c) What is the probability of all three of the events occurring simultaneously?

d) What is the probability of at least one of E, LL, or G occurring?

Q29. Over the course of two baseball seasons, player A obtained 126 hits in 500 at-bats in Season 1, and 90 hits in 300 at-bats in Season 2; player B, on the other hand, obtained 75 hits in 300 at-bats in Season 1, and 145 hits in 500 at-bats in Season 2. A player's batting average is the number of hits they obtain divided by the number of at-bats.

a) Which player has the best batting average in Season 1? In Season 2?

b) Which player has the best batting average over the 2-year period?

c) What is happening here?

Q30. A stranger comes to you and shows you what appears to be a normal coin, with two distinct sides: Heads (H) and Tails (T). They flip the coin 4 times and record the following sequence of tosses: $HHHH$.

a) What is the probability of obtaining this specific sequence of tosses? What assumptions do you make along the way in order to compute the probability? What is the probability that the next toss will be a T.

b) The stranger offers you a bet: they will toss the coin another time; if the toss is T, they give you $100\$$, but if it is H, you give them $10\$$. Would you accept the bet (if you are not morally opposed to gambling)?

c) Now the stranger tosses the coin 60 times and records $60 \times H$ in a row: $H \cdots H$. They offer you the same bet. Do you accept it?

d) What if they offered $1000\$$ instead? $1,000,000\$$?

Q31. The sample space of a random experiment is $\{a, b, c, d, e, f\}$ and each outcome is equally likely. A random variable is defined as follows

outcome	a	b	c	d	e	f
X	0	0	1.5	1.5	2	3

Determine the probability mass function of X. Determine the following probabilities:

a)$P(X = 1.5)$ b)$P(0.5 < X < 2.7)$ c)$P(X > 3)$

d)$P(0 \leq X < 2)$ e)$P(X = 0 \text{ or } 2)$

Q32. Determine the mean and the variance of the random variable defined in **Q1**.

Q33. We say that X has **uniform distribution** on a set of values $\{X_1, \ldots, X_k\}$ if

$$P(X = X_i) = \frac{1}{k}, \qquad i = 1, \ldots, k.$$

The thickness measurements of a coating process are **uniformly distributed** with values 0.15, 0.16, 0.17, 0.18, 0.19. Determine the mean and variance of the thickness measurements.

Is this result compatible with a uniform distribution?

Q34. Samples of rejuvenated mitochondria are mutated in 1% of cases. Suppose 15 samples are studied and that they can be considered to be independent (from a mutation standpoint). Determine the following probabilities:

a) no samples are mutated;
b) at most one sample is mutated, and
c) more than half the samples are mutated.

Use the following CDF table for the $\mathcal{B}(n, p)$, with $n = 15$ and $p = 0.99$:

r	0	1	2	3	4	5	6	7
$P(X \leq r)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
r	8	9	10	11	12	13	14	15
$P(X \leq r)$	0.0000	0.0000	0.0000	0.0000	0.0004	0.0096	0.1399	1.0000

Q35. Samples of 20 parts from a metal punching process are selected every hour. Typically, 1% of the parts require re-work. Let X denote the number of parts in the sample that require re-work. A process problem is suspected if X exceeds its mean by more than three standard deviations.

a) What is the probability that there is a process problem?

b) If the re-work percentage increases to 4%, what is the probability that X exceeds 1?

c) If the re-work percentage increases to 4%, what is the probability that X exceeds 1 in at least one of the next five sampling hours?

Q36. In a clinical study, volunteers are tested for a gene that has been found to increase the risk for a particular disease. The probability that the person carries a gene is 0.1.

a) What is the probability that 4 or more people will have to be tested in order to detect 1 person with the gene?

b) How many people are expected to be tested in order to detect 1 person with the gene?

c) How many people are expected to be tested before 2 with the gene are detected?

Q37. The number of failures of a testing instrument from contaminated particles on the product is a Poisson random variable with a mean of 0.02 failure per hour.

a) What is the probability that the instrument does not fail in an $8-$hour shift?

b) What is the probability of at least 1 failure in a $24-$hour day?

Q38. Use `R` to generate a sample from a binomial distribution and from a Poisson distribution (select parameters as you wish).

Use `R` to compute the sample means and sample variances. Compare these values to population means and population variances.

Q39. A container of 100 light bulbs contains 5 bad bulbs. We draw 10 bulbs without replacement. Find the probability of drawing at least 1 defective bulb.

a) 0.4164 b) 0.584 c) 0.1 d) 0.9 e) none of the preceding

Q40. Let X be a disrete random variable with range $\{0, 1, 2\}$ and probability mass function (p.m.f.) given by $f(0) = 0.5$, $f(1) = 0.3$, and $f(2) = 0.2$. The expected value and variance of X are, respectively,

a)0.7, 0.61 b)0.7, 1.1 c)0.5, 0.61 d)0.5, 1.1 e)none of the preceding

Q41. A factory employs several thousand workers, of whom 30% are not from an English-speaking background. If 15 members of the union executive committee were chosen from the workers at random, evaluate the probability that exactly 3 members of the committee are not from an English-speaking background.

a) 0.17 b) 0.83 c) 0.98 d) 0.51 e) none of the preceding

Use the following CDF table for the $\mathcal{B}(n,p)$, with $n = 15$ and $p = 0.30$ if needed:

r	0	1	2	3	4	5	6	7
$P(X \leq r)$	0.0047	0.0353	0.1268	0.2969	0.5155	0.7216	0.8689	0.9500
r	8	9	10	11	12	13	14	15
$P(X \leq r)$	0.9848	0.9963	0.9993	0.9999	1.0000	1.0000	1.0000	1.0000

Q42. Assuming the context of **Q11**, what is the probability that a majority of the committee members do not come from an English-speaking background?

Q43. In a video game, a player is confronted with a series of opponents and has an 80% probability of defeating each one. Success with any opponent (that is, defeating the opponent) is independent of previous encounters. The player continues until defeated. What is the probability that the player encounters at least three opponents?

a) 0.8 b) 0.64 c) 0.5 d) 0.36 e) none of the preceding

Q44. Assuming the context of **Q13**, how many encounters is the player expected to have?

a) 5 b) 4 c) 8 d) 10 e) none of the preceding

Q45. From past experience it is known that 3% of accounts in a large accounting company are in error. The probability that exactly 5 accounts are audited before an account in error is found, is:

a)0.242 b)0.011 c)0.030 d)0.026 e)none of the preceding

Q46. A receptionist receives on average 2 phone calls per minute. Assume that the number of calls can be modeled using a Poisson random variable. What is the probability that he does not receive a call within a $3-$minute interval?

a)e^{-2} b)$e^{-1/2}$ c)e^{-6} d)e^{-1} e)none of the preceding

Q47. Consider a random variable X with probability density function (p.d.f.) given by

$$f(x) = \begin{cases} 0 & \text{if } x \le -1 \\ 0.75(1-x^2) & \text{if } -1 \le x < 1 \\ 0 & \text{if } x \ge 1 \end{cases}$$

What is the expected value and the standard deviation of X?

a) $0,\ 3$ b) $0,\ 0.447$ c) $1,\ 0.2$ d) $1,\ 3$ e) none of the preceding

Q48. A random variable X has a cumulative distribution function (c.d.f.)

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x/2 & \text{if } 0 < x < 2 \\ 1 & \text{if } x \geq 2 \end{cases}$$

What is the mean value of X?

a) 1 b) 2 c) 0 d) 0.5 e) none of the preceding

Q49. Let X be a random variable with p.d.f. given by $f(x) = \frac{3}{2}x^2$ for $-1 \leq x \leq 1$, and $f(x) = 0$ otherwise. Find $P(X^2 \leq 0.25)$.

a)0.250 b)0.125 c)0.500 d)0.061 e)none of the preceding

Q50. In the inspection of tin plate produced by a continuous electrolytic process, 0.2 imperfections are spotted per minute, on average. Find the probability of spotting at least 2 imperfections in 5 minutes. Assume that we can model the occurrences of imperfections as a Poisson process.

a) 0.736 b) 0.264 c) 0.632 d) 0.368 e) none of the preceding

Q51. If $X \sim \mathcal{N}(0, 4)$, the value of $P(|X| \geq 2.2)$ is (using the normal table):

a)0.2321 b)0.8438 c)0.2527 d)0.2713 e)0.7286 f) none of the preceding

Q52. If $X \sim \mathcal{N}(10, 1)$, the value of k such that $P(X \leq k) = 0.701944$ is closest to

a)0.59 b)0.30 c)0.53 d)10.53 e)10.30 f)10.59

Q53. The time it takes a supercomputer to perform a task is normally distributed with mean 10 milliseconds and standard deviation 4 milliseconds. What is the probability that it takes more than 18.2 milliseconds to perform the task? (use the normal table or `R`).

a) 0.9798 b) 0.8456 c) 0.0202 d) 0.2236 e) 0.5456 f) none of the preceding

Q54. Roll a $4-$sided die twice, and let X equal the larger of the two outcomes if they are different and the common value if they are the same. Find the p.m.f. and the c.d.f. of X.

Q55. Compute the mean and the variance of X as defined in **Q24**, as well as $\mathrm{E}[X(5-X)]$.

Q56. In 80% of cases when a basketball player attempts a free throw, they are successful. Assume that each of the free throw attempts are independent. Let X be the minimum number of attempts in order to succeed 10 times. Find the p.m.f. of X and the probability that $X = 12$.

Q57. Let X be the minimum number of independent trials (each with probability of success p) that are needed to observe r successes. The p.m.f. of X is

$$f(x) = P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-1}, \quad x = r, r+1, \ldots$$

The mean and variance of X are $\mathrm{E}[X] = \frac{r}{p}$ and $\mathrm{Var}[X] = \frac{r(1-p)}{p^2}$. Compute the mean minimum number of independent free throw attempts required to observe 10 successful free throws if the probability of success at the free thrown line is 80%. What about the standard deviation of X?

Q58. If $n \geq 20$ and $p \leq 0.05$, it can be shown that the binomial distribution with n trials and an independent probability of success p can be approximated by a Poisson distribution with parameter $\lambda = np$. This is called the **Poisson approximation**:

$$\frac{(np)^x e^{-np}}{x!} \approx \binom{n}{x} p^x (1-p)^{n-x}.$$

A manufacturer of light bulbs knows that 2% of its bulbs are defective. What is the probability that a box of 100 bulbs contains exactly at most 3 defective bulbs? Use the Poisson approximation to estimate the probability.

Q59. Consider a discrete random variable X which has a uniform distribution over the first positive m integers, i.e.

$$f(x) = P(X = x) = \frac{1}{m}, \quad x = 1, \ldots, m,$$

and $f(x) = 0$ otherwise. Compute the mean and the variance of X. For what values of m is $\mathrm{E}[X] > \mathrm{Var}[X]$?

Q60. Let X be a random variable. What is the value of b (where b is not a function of X) which minimizes $\mathrm{E}[(X-b)^2]$?

Q61. An experiment consists in selecting a bowl, and then drawing a ball from that bowl. Bowl B_1 contains two red balls and four white balls; bowl B_2 contains one red ball and two white balls; and bowl B_3 contains five red balls and four white balls. The probabilities for selecting the bowls are not uniform: $P(B_1) = 1/3$, $P(B_2) = 1/6$, and $P(B_3) = 1/2$, respectively.

a) What is the probability of drawing a red ball $P(R)$?

b) If the experiment is conducted and a red ball is drawn, what is the probability that the ball was drawn from bowl B_1? B_2? B_3?

Q62. The time to reaction to a visual signal follows a normal distribution with mean 0.5 seconds and standard deviation 0.035 seconds.

a) What is the probability that time to react exceeds 1 second?

b) What is the probability that time to react is between 0.4 and 0.5 seconds?

c) What is the time to reaction that is exceeded with probability of 0.9?

Q63. Suppose that the random variable X has the following cumulative distribution function:

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ x^3, & 0 \leq x \leq 1 \\ 1, & x \geq 1. \end{cases}$$

a) Compute $P(X > 0.5)$.

b) Compute $P(0.2 < X < 0.8)$.

c) Find the probability density function of X.

d) Find $\mathrm{E}[X]$ and $\mathrm{Var}[X]$.

Q64. Assume that arrivals of small aircrafts at an airport can be modeled by a Poisson random variable with an average of 1 aircraft per hour.

a) What is the probability that more than 3 aircrafts arrive within an hour?

b) Consider 15 consecutive and disjoint $1-$hour intervals. What is the probability that in none of these intervals we have more than 3 aircraft arrivals?

c) What is the probability that exactly 3 aircrafts arrive within 2 hours?

Q65. Refer to the situation described in **Q3**.

d) What is the length of the interval such that the probability of having no arrival within this interval is 0.1?

e) What is the probability that one has to wait at least 3 hours for the arrival of 3 aircrafts?

f) What is the mean and variance of the waiting time for 3 aircrafts?

Q66. Assume that X is normally distributed with mean 10 and standard deviation 3. In each case, find the value x such that:

a) $P(X > x) = 0.5$

b) $P(X > x) = 0.95$

c) $P(x < X < 10) = 0.2$

d) $P(-x < X - 10 < x) = 0.95$

e) $P(-x < X - 10 < x) = 0.99$

Q67. Let $X \sim \text{Exp}(\lambda)$ with mean 10. What is $P(X > 30|X > 10)$ equal to?

a)$1 - \exp(-2)$ b)$\exp(-2)$ c)$\exp(-3)$

d)$1/10$ e)$\exp(-200)$ f) none of the preceding

Q68. Let X denote a number of failures of a particular machine within a month. Its probability mass function is given by

x	0	1	2	3	4	5
$P(X=x)$	0.17	0.23	0.19	0.13	0.08	0.2

The probability that there are fewer than 3 failures within a month, and the expected number of failures within a month are, respectively,

a)$0.28; 2.50$ b)$0.72; 2.32$ c)$0.59; 2.32$

d)$0.80; 2.50$ e)none of the preceding

Q69. A company's warranty document states that the probability that a new swimming pool requires some repairs within the first year is 20%. What is the probability, that the sixth sold pool is the first one which requires some repairs within the first year?

a)0.6068 b)0.3932 c)0.9345 d)0.0655 e)none of the preceding

Q70. In a group of ten students, each student has a probability of 0.7 of passing the exam. What is the probability that exactly 7 of them will pass an exam?

a) 0.9829 b) 0.2668 c) 0.0480 d) 0.9520 e) none of the preceding

Q71. Two companies A and B consider making an offer for road construction. The company A makes the submission. The probability that B submits the proposal is $1/3$. If B does not submit the proposal, the probability that A gets the job is $3/5$. If B submits the proposal, the probability that A gets the job is $1/3$. What is the probability that A will get the job?

a) 0.6667 b) 0.5111 c) 0.7500 d) 0.3333 e) none of the preceding

Q72. In a box of 50 fuses there are 8 defective ones. We choose 5 fuses randomly (without replacement). What is the probability that all 5 fuses are not defective?

a) 0.4015 b) 0.84 c) 0.3725 d) 0.4275 e) none of the preceding

Q73. Consider a random variable X with the following probability density function:

$$f(x) = \begin{cases} 0 & \text{if } x \leq -1 \\ \frac{3}{4}(1 - x^2) & \text{if } -1 < x < 1 \\ 0 & \text{if } x \geq 1 \end{cases}$$

The value of $P(X \leq 0.5)$ is

a) $11/32$ b) $27/32$ c) $16/32$ d) 1 e) none of the preceding

Q74. A receptionist receives on average 2 phone calls per minute. If the number of calls follows a Poisson process, what is the probability that the waiting time for call will be greater than 1 minute?

a) $e^{-1/15}$ b) $e^{-1/30}$ c) e^{-2} d) e^{-1} e) none of the preceding

Q75. A company manufactures hockey pucks. It is known that their weight is normally distributed with mean 1 and standard deviation 0.05. The pucks used by the NHL must weigh between 0.9 and 1.1. What is the probability that a randomly chosen puck can be used by NHL?

a) 1 b) 0.9545 c) 0.4560 d) 0.9772 e) none of the preceding

Q76. Consider the following dataset:

`12 14 6 10 1 20 4 8`

The median and the first quartile of the dataset are, respectively:

a) $9,\ 5$ b) $5.5,\ 6$ c) $10,\ 5$ d) $5,\ 10$ e) none of the preceding

Q77. Let X denote a number of failures of a particular machine within a month. Its probability mass function is given by

x	0	1	2	3	4	5
$P(X = x)$	0.17	0.23	0.19	0.13	0.08	0.2

- The probability that there are less than 3 failures within a month, and
- the expected number of failures within a month

are, respectively:

a) $0.28;\ 2.50$ b) $0.72;\ 2.32$ c) $0.59;\ 2.32$ d) $0.80;\ 2.50$ e) none of the preceding

Q78. Consider the following `R` output:

```
> pbinom(15,100,0.25)
[1] 0.01108327
> pbinom(17,100,0.25)
[1] 0.03762626
> pbinom(31,100,0.25)
[1] 0.9306511
```

```
> pbinom(16,100,0.25)
[1] 0.02111062
> pbinom(30,100,0.25)
[1] 0.8962128
> pbinom(32,100,0.25)
[1] 0.9554037
```

Let X be a binomial random variable with $n = 100$ and $p = 0.25$. Using the `R` output above, calculate $P(16 \leq X \leq 31)$.

a)0.9196 b)0.9095 c)0.9348 d)0.9443 e)none of the preceding

Q79. Suppose that samples of size $n = 25$ are selected at random from a normal population with mean 100 and standard deviation 10. What is the probability that sample mean falls in the interval

$$(\mu_{\overline{X}} - 1.8\sigma_{\overline{X}}, \mu_{\overline{X}} + 1.0\sigma_{\overline{X}})?$$

Q80. The compressive strength of concrete is normally distributed with mean $\mu = 2500$ and standard deviation $\sigma = 50$. A random sample of size 5 is taken. What is the standard error of the sample mean?

Q81. Suppose that $X_1 \sim \mathcal{N}(3, 4)$ and $X_2 \sim \mathcal{N}N(3, 45)$. Given that X_1 and X_2 are independent random variables, what is a good approximation to $P(X_1 + X_2 > 9.5)$?

a) 0.3085 b) 0.6915 c) 0.5279 d) 0.4271 e) none of the preceding

Q82. The amount of time that a customer spends waiting at an airport check-in counter is a random variable with mean $\mu = 8.2$ minutes and standard deviation $\sigma = 1.5$ minutes. Suppose that a random sample of $n = 49$ customers is taken. Compute the approximate probability that the average waiting time for these customers is:

a) Less than 10 min.

b) Between 5 and 10 min.

c) Less than 6 min.

Q83. A random sample of size $n_1 = 16$ is selected from a normal population with a mean of 75 and standard deviation of 8. A second random sample of size $n_2 = 9$ is taken independently from another normal population with mean 70 and standard deviation of 12. Let $\overline{X}_1$ and $\overline{X}_2$ be the two sample means. Find

a) The probability that $\overline{X}_1 - \overline{X}_2$ exceeds 4.

b) The probability that $3.5 < \overline{X}_1 - \overline{X}_2 < 5.5$.

Q84. Discuss the normality of the following dataset:

170,295,200,165,140,190,195,142,138,148,110,140,103,176,125,
126,204,196,98,123,124,152,177,168,175,186,140,147,174,155,195

Q85. Using `R`, illustrate the central limit theorem by generating $M = 300$ samples of size $n = 30$ from:

- a normal random variable with mean 10 and variance 0.75;
- a binomial random variable with 3 trials and probability of success 0.3.

Repeat the same procedure for samples of size $n = 200$. What do you observe?

Hint: In each case, assess the normality using a histogram and a QQ plot.

Q86. Suppose that the weight in pounds of a North American adult can be represented by a normal random variable with mean 150 lbs and variance 900 lbs^2. An elevator containing a sign "Maximum 12 people" can safely carry 2000 lbs. The probability that 12 North American adults will not overload the elevator is closest to

a) 0.9729 b) 0.4501 c) 0.0271 d) 0.0001 e) 1.3 f) none of the preceding

Q87. Let $X_1, \cdots, X_{50}$ be an independent random sample from a Poisson distribution with mean 1. Set $Y = X_1 + \cdots + X_{50}$. The approximate probability $P(48 \leq Y \leq 52)$ is closest to:

a) 0.6368 b) 0.4534 c) 0.2227 d) 0.9988 e) 0.5000 f) none of the preceding

Q88. A new type of electronic flash for cameras will last an average of 5000 hours with a standard deviation of 500 hours. A quality control engineer intends to select a random sample of 100 of these flashes and use them until they fail. What is the probability that the mean life time of the sample of 100 flashes will be less than 4928 hours?

a)0.0749 b)0.9251 c)0.0002 d)0.4532 e)none of the preceding

Q89. A manufacturer of fluoride toothpaste regularly measures the concentration of of fluoride in the toothpaste to make sure that it is within the specifications of $0.85 - 1.10$ mg/g. The table on the next page lists 100 such measurements. Build a relative frequency histogram of the data (a histogram with area $= 1$).

Table 6.1-3 Concentrations of fluoride in mg/g in toothpaste

0.98	0.92	0.89	0.90	0.94	0.99	0.86	0.85	1.06	1.01
1.03	0.85	0.95	0.90	1.03	0.87	1.02	0.88	0.92	0.88
0.88	0.90	0.98	0.96	0.98	0.93	0.98	0.92	1.00	0.95
0.88	0.90	1.01	0.98	0.85	0.91	0.95	1.01	0.88	0.89
0.99	0.95	0.90	0.88	0.92	0.89	0.90	0.95	0.93	0.96
0.93	0.91	0.92	0.86	0.87	0.91	0.89	0.93	0.93	0.95
0.92	0.88	0.87	0.98	0.98	0.91	0.93	1.00	0.90	0.93
0.89	0.97	0.98	0.91	0.88	0.89	1.00	0.93	0.92	0.97
0.97	0.91	0.85	0.92	0.87	0.86	0.91	0.92	0.95	0.97
0.88	1.05	0.91	0.89	0.92	0.94	0.90	1.00	0.90	0.93

Q90. Use the data from **Q89**.

a) Compute the data's mean $\overline{x}$ and it's standard deviation s_x (use a computer program, for goodness' sake!)

b) Using the frequency table of fluoride concentrations (Table 6.1-4), you can also approximate the mean and variance. Let u_i be the **class mark** for each of the histogram's 8 classes (the midpoint along the rectangles' widths), n be the total number of observations, and k be the number of classes. Then

$$\overline{u} = \frac{1}{n}\sum_{i=1}^{k} f_i u_i \quad \text{and} \quad s_u^2 = \frac{1}{n-1} f_i (u_i - \overline{u})^2.$$

Compute $\overline{u}$ and s_u. How do they compare with $\overline{x}$ and s_x?

Q91. Use the data from **Q89**.

a) Provide a the $5-$number summary of the data $(q_0, q_1, q_2, q_3, q_4)$, as well as the interquartile range IQR.

b) Display the $5-$number summary as a boxplot chart.

Q92. Use the data from **Q89**. Compute the **midrange** $\frac{1}{2}(Q_0 + Q_4)$, the **trimean** $\frac{1}{4}(Q_1 + 2Q_2 + Q_3)$, and the **range** $Q_4 - Q_0$ for the fluoride data.

Q93. A new cure has been developed for a certain type of cement that should change its mean compressive strength. It is known that the standard deviation of the compressive strength is 130 kg/cm^2 and that we may assume that it follows a normal distribution. 9 chunks of cement have been tested and the observed sample mean is $\overline{X} = 4970$. Find the 95% confidence interval for the mean of the compressive strength.

a)$[4858.37, 5081.63]$ b)$[4885.07, 5054.93]$ c)$[4858.37, 5054.93]$

d)$[4944.52, 4995.48]$ e)none of the preceding

Q94. Consider the same set-up as in **Q93**, but now 100 chunks of cement have been tested and the observed sample mean is $\overline{X} = 4970$. Find the 95% confidence interval for the mean of the compressive strength.

a)$[4858.37, 5081.63]$ b)$[4885.07, 5054.93]$ c)$[4858.37, 5054.93]$

d)$[4944.52, 4995.48]$ e)none of the preceding

Q95. Consider the same set-up as in **Q93**, but now we do not know the standard deviation of the normal distribution. 9 chunks of cement have been tested, and the measurements are

$$5001, 4945, 5008, 5018, 4991, 4990, 4968, 5020, 5003.$$

Find the 95% confidence interval for the mean of the compressive strength.

a) $[4858.37, 5081.63]$ b) $[4885.07, 5054.93]$ c) $[4858.37, 5054.93]$

d) $[4944.52, 4995.48]$ e) none of the preceding

Q96. A steel bar is measured with a device which a known precision of $\sigma = 0.5$mm. Suppose we want to estimate the mean measurement with an error of at most 0.2mm at a level of significance $\alpha = 0.05$. What sample size is required? Assume normality.

a)25 b)24 c)6 d)7 e)none of the preceding

Q97. In a random sample of 1000 houses in the city, it is found that 228 are heated by oil. Find a 99% C.I. for the proportion of homes in the city that are heated by oil.

a)$[0.202, 0.254]$ b)$[0.197, 0.259]$ c)$[0.194, 0.262]$

d)$[0.185, 0.247]$ e)none of the preceding

Q98. Past experience indicates that the breaking strength of yarn used in manufacturing drapery material is normally distributed and that $\sigma = 2$ psi. A random sample of 15 specimens is tested and the average breaking strength is found to be $\overline{x} = 97.5$ psi.

a) Find a 95% confidence interval on the true mean breaking strength.

b) Find a 99% confidence interval on the true mean breaking strength.

Q99. The diameter holes for a cable harness follow a normal distribution with $\sigma = 0.01$ inch. For a sample of size 10, the average diameter is 1.5045 inches.

a) Find a 99% confidence interval on the mean hole diameter.

b) Repeat this for $n = 100$.

Q100. A journal article describes the effect of delamination on the natural frequency of beams made from composite laminates. The observations are as follows:

$$230.66, 233.05, 232.58, 229.48, 232.58, 235.22.$$

Assuming that the population is normal, find a 95% confidence interval on the mean natural frequency.

Q101. A textile fiber manufacturer is investigating a new drapery yarn, which the company claims has a mean thread elongation of $\mu = 12$ kilograms with standard deviation of $\sigma = 0.5$ kilograms.

a) What should be the sample size so that with probability 0.95 we will estimate the mean thread elongation with error at most 0.15 kg?

b) What should be the sample size so that with probability 0.95 we will estimate the mean thread elongation with error at most 0.05 kg?

Q102. The brightness of television picture tube can be evaluated by measuring the amount of current required to achieve a particular brightness level. An engineer thinks that one has to use 300 microamps of current to achieve the required brightness level. A sample of size $n = 20$ has been taken to verify the engineer's hypotheses.

a) Formulate the null and the alternative hypotheses. Use a two-sided test alternative.

b) For the sample of size $n = 20$ we obtain $\overline{x} = 319.2$ and $s = 18.6$. Test the hypotheses from part a) with $\alpha = 5\%$ by computing a critical region. Calculate the p-value.

c) Use the data from part b) to construct a 95% confidence interval for the mean required current.

Q103. We say that a particular production process is **stable** if it produces at most 2% defective items. Let p be the true proportion of defective items.

a) We sample $n = 200$ items at random and consider hypotheses testing about p. Formulate null and alternative hypotheses.

b) What is your conclusion of the above test, if one observes 3 defective items out of 200? Note: you have to choose an appropriate level α.

Q104. Ten engineers' knowledge of basic statistical concepts was measured on a scale of $0 - 100$, before and after a short course in statistical quality control. The result are as follows:

Engineer	1	2	3	4	5	6	7	8	9	10
Before X_{1i}	43	82	77	39	51	66	55	61	79	43
After X_{2i}	51	84	74	48	53	61	59	75	82	53

Let μ_1 and μ_2 be the mean mean score before and after the course. Perform the test $H_0 : \mu_1 = \mu_2$ against $H_A : \mu_1 < \mu_2$. Use $\alpha = 0.05$.

105. 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 B's 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 supplier's A rods, respectively, and the same parameters are 657 MPa and 3 MPa for supplier B's rods. Perform the test $H_0 : \mu_A = \mu_B$ against $\mu_A \neq \mu_B$. Use $\alpha = 0.05$. Assume that the variances are equal but unknown.

106. The deflection temperature under load for two different types of plastic pipe is being investigated. Two random samples of 15 pipe specimens are tested, and the deflection temperatures observed are as follows:

Type 1: $206, 188, 205, 187, 194, 193, 207, 185, 189, 213, 192, 210, 194, 178, 205.$

Type 2: $177, 197, 206, 201, 180, 176, 185, 200, 197, 192, 198, 188, 189, 203, 192.$

Does the data support the claim that the deflection temperature under load for type 1 pipes exceeds that of type 2? Calculate the p-value, using $\alpha = 0.05$, and state your conclusion.

Q107. It is claimed that 15% of a certain population is left-handed, but a researcher doubts this claim. They decide to randomly sample 200 people and use the anticipated small number to provide evidence against the claim of 15%. Suppose 22 of the 200 are left-handed. Compute the $p-$value associated with the hypothesis (assuming a binomial distribution), and provide an interpretation.

Q108. A child psychologist believes that nursery school attendance improves children's social perceptiveness (SP). They use 8 pairs of twins, randomly choosing one to attend nursery school and the other to stay at home, and then obtains scores for all 16. In 6 of the 8 pairs, the twin attending nursery school scored better on the SP test. Compute the $p-$value associated with the hypothesis (assuming a binomial distribution), and provide an interpretation.

Q109. It is claimed that the breaking strength of yarn used in manufacturing drapery material is normally distributed with mean 97 and $\sigma = 2$ psi. A random sample of nine specimens is tested and the average breaking strength is found to be $\overline{X} = 98$ psi. Formulate a test for this situation. Should it be $1-$sided or $2-$sided? What value of α should you use? What conclusion does the test and the sample yield?

Q110. A civil engineer is analyzing the compressive strength of concrete. It is claimed that its mean is 80 and variance is known to be 2. A random sample of size 60 yields the sample mean 59. Formulate a test for this situation. Should it be $1-$sided or $2-$sided? What value of α should you use? What conclusion does the test and the sample yield?

Q111. The sugar content of the syrup in canned peaches is claimed to be normally distributed with mean 10 and variance 2. A random sample of $n = 10$ cans yields a sample mean 11. Another random sample of $n = 10$ cans yields a sample mean 9. Formulate a test for this situation. Should it be $1-$sided or $2-$sided? What value of α should you use? What conclusion does the test and the sample yield?

Q112. A certain power supply is stated to provide a constant voltage output of 10kV. Ten measurements are taken and yield the sample mean of 11kV. Formulate a test for this situation. Should it be $1-$sided or $2-$sided? What value of α should you use? What conclusion does the test and the sample yield?

Q113. The mean water temperature downstream from a power water plant cooling tower discharge pipe should be no more than 100F. Past experience has indicated that that the standard deviation is 2F. The water temperature is measured on nine randomly chosen days, and the average temperature is found to be 98F. Formulate a test for this situation. Should it be $1-$sided or $2-$sided? What value of α should you use? What conclusion does the test and the sample yield?

Q114. We are interested in the mean burning rate of a solid propellant used to power aircrew escape systems. We want to determine whether or not the mean burning rate is 50 cm/second. A sample of 10 specimens is tested and we observe $\overline{X} = 48.5$. Assume normality with $\sigma = 2.5$.

Q115. Ten individuals have participated in a diet modification program to stimulate weight loss. Their weight both before and after participation in the program is shown below:

Before	$195, 213, 247, 201, 187, 210, 215, 246, 294, 310$
After	$187, 195, 221, 190, 175, 197, 199, 221, 278, 285$

Is there evidence to support the claim that this particular diet-modification program is effective in producing mean weight reduction? Use $\alpha = 0.05$. Compute the associated $p-$value.

Q116. We want to test the hypothesis that the average content of containers of a particular lubricant equals 10L against the two-sided alternative. The contents of a random sample of 10 containers are

10.2	9.7	10.1	10.3	10.1
9.8	9.9	10.4	10.3	9.5

Find the $p-$value of this two-sided test. Assume that the distribution of contents is normal. Note that $\sum_{i=1}^{10} x_i^2 = 1006.79$, if x_i represent the measurements.

a)$0.05 < p < 0.10$ b)$0.10 < p < 0.20$ c)$0.25 < p < 0.40$

d)$0.50 < p < 0.80$ e)none of the preceding

Q117. An engineer measures the weight of $n = 25$ pieces of steel, which follows a normal distribution with variance 16. The average weight for the sample is $\overline{X} = 6$. They want to test for $H_0 : \mu = 5$ against $H_1 : \mu > 5$. What is the $p-$value for the test?

a) 0.05000 b) 0.10565 c) 0.89435 d) 1.0000 e) none of the preceding

Q118. The thickness of a plastic film (in mm) on a substrate material is thought to be influenced by the temperature at which the coating is applied. A completely randomized experiment is carried out. 11 substrates are coated at 125F, resulting in a sample mean coating thickness of $\overline{x}_1 = 103.5$ and a sample standard deviation of $s_1 = 10.2$. Another 11 substrates are coated at 150F, for which $\overline{x}_2 = 99.7$ and $s_2 = 11.7$ are observed. We want to test equality of means against the two-sided alternative. The value of the appropriate test statistics and the decision are (for $\alpha = 0.05$):

a)0.81; Reject H_0.

b)0.81; Do not reject H_0.

c)1.81; Reject H_0.

d)1.81; Do not reject H_0.

e)none of the preceding

Note: assume that population variances are unknown but equal.

Q119. The following output was produced with `t.test` command in `R`.

```
One Sample t-test
data:  x
t = 2.0128, df = 99, p-value = 0.02342
alternative hypothesis: true mean is greater than 0
```

Based on this output, which statement is correct?

a) If the type I error is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;
b) If the type I error is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu \neq 0$;
c) If the type I error is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;
d) If the type I error is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu < 0$;
e) Type I error is 0.02342.

Q120. A pharmaceutical company claims that a drug decreases a blood pressure. A physician doubts this claim. They test 10 patients and records results before and after the drug treatment:

```
> Before=c(140,135,122,150,126,138,141,155,128,130)
> After=c(135,136,120,148,122,136,140,153,120,128)
```

At the `R` command prompt, they type:

```
> test.t(Before,After,alternative="greater")
    data:  Before and After
    t = 0.5499, p-value = 0.2946
    alternative hypothesis: true difference in means is greater than 0
    sample estimates: mean of x mean of y
        136.5     133.8
```

Their assistant claims that the command should instead be:

```
> test.t(Before,After,paired=TRUE,alternative="greater")
```

```
data: Before and After t = 3.4825, df = 9, p-value = 0.003456
alternative hypothesis: true difference in means is greater than 0
sample estimates: mean of the differences
    2.7
```

Which answer is best?

a) The assistant uses the correct command. There is not enough evidence to justify that the new drug decreases blood pressure;
b) The assistant uses the correct command. There is enough evidence to justify that the new drug decreases blood pressure for any reasonable choice of α;
c) The physician uses the correct command. There is not enough evidence to justify that the new drug decreases blood pressure;
d) The physician uses the correct command. There is enough evidence to justify that the new drug decreases blood pressure for any reasonable choice of α;
e) Nobody is correct, $t-$tests should not be used here.

Q121. A company claims that the mean deflection of a piece of steel which is 10ft long is equal to 0.012ft. A buyer suspects that it is bigger than 0.012ft. The following data x_i has been collected:

```
0.0132 0.0138 0.0108 0.0126 0.0136 0.0112 0.0124 0.0116 0.0127 0.0131
```

Assuming normality and that $\sum_{i=1}^{10} x_i^2 = 0.0016$, what are the $p-$value for the appropriate one-sided test and the corresponding decision?

a)$p \in (0.05, 0.1)$ and reject H_0 at $\alpha = 0.05$.

b)$p \in (0.05, 0.1)$ and do not reject H_0 at $\alpha = 0.05$.

c)$p \in (0.1, 0.25)$ and reject H_0 at $\alpha = 0.05$.

d)$p \in (0.1, 0.25)$ and do not reject H_0 at $\alpha = 0.05$.

e)none of the preceding

Q122. 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; 11 pieces of type B sandpaper were subjected to the same treatment. We have the following observations:

```
xA 27 26 24 29 30 26 27 23 28 27
xB 24 23 22 27 24 21 24 25 24 23 20
```

Note that $\sum x_{A,i} = 267$, $\sum x_{B,i} = 257$, $\sum x_{A,i}^2 = 7169$, $\sum x_{B,i}^2 = 6041$. Assuming normality and equality of variances in abrasive wear for A and B, we want to test for equality of mean abrasive wear for A and B. The appropriate $p-$value is

a)$p < 0.01$ b)$p > 0.2$ c)$p \in (0.01, 0.05)$

d)$p \in (0.1, 0.2)$ e)$p \in (0.05, 0.1)$ f) none of the preceding

Q123. 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
```

Based on this output, which statement is correct?

a) If the type I error is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;
b) If the type I error is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu \neq 0$;
c) If the type I error is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;
d) If the type I error is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu < 0$;
e) None of the preceding.

Q124. Consider a sample $\{X_1, \ldots, X_{10}\}$ from a normal population $X_i \sim \mathcal{N}(4, 9)$. Denote by $\overline{X}$ and S^2 the sample mean and the sample variance, respectively. Find c such that

$$P\left(\frac{\overline{X} - 4}{S/\sqrt{10}} \leq c\right) = 0.99$$

a) 1.833 b) 2.326 c) 1.645 d) 2.821 e) none of the preceding

Q125. Consider the following dataset:

```
2.6  3.7  0.8  9.6  5.8 -0.8  0.7 0.6
4.8  1.2  3.3  5.0  3.7  0.1 -3.1 0.3
```

The median and the interquartile range of the sample are, respectively:

a) $2.4, 3.3$ b) $1.9, 3.8$ c) $1.9, 1.8$ d) $2.9, 12.2$ e) none of the preceding

Q126. An article in *Computers and Electrical Engineering* considered the speed-up of cellular neural networks (CNN) for a parallel general-purpose computing architecture. Various speed-ups are observed:

```
3.77  3.35  4.21  4.03  4.03  4.63
4.63  4.13  4.39  4.84  4.26  4.60
```

Assume that the population is normally distributed. The 99% C.I. for the mean speed-up is:

a) $[4.155, 4.323]$ b) $[3.863, 4.615]$ c) $[4.040, 4.438]$

d) $[3.77, 4.60]$ e) none of the preceding

Q127. An engineer measures the weight of $n = 25$ pieces of steel, which follows a normal distribution with variance 16. The average observed weight for the sample is $\overline{x} = 6$. The two-sided 95% C.I. for the mean μ is:

a)$[-0.272, 12.272]$ b)$[4.432, 7.568]$ c)$[3.250, 8.750]$

d)$[4.120, 7.522]$ e)none of the preceding

Q128. Assume that random variables $\{X_1, \ldots, X_8\}$ follow a normal distribution with mean 2 and variance 24. Independently, assume that random variables $\{Y_1, \ldots, X_{16}\}$ follow a normal distribution with mean 1 and variance 16. Let $\overline{X}$ and $\overline{Y}$ be the corresponding sample means. Then $P(\overline{X} + \overline{Y} > 4)$ is:

a)0.7721 b)0.30855 c)0.69165 d)0.9883 e)none of the preceding

Q129. A medical team wants to test whether a particular drug decreases diastolic blood pressure. Nine people have been tested. The team measured blood pressure before (X) and after (Y) applying the drug. The corresponding means were $\overline{X} = 91$, $\overline{Y} = 87$. The sample variance of the differences was $S_D^2 = 25$. The $p-$value for the appropriate one-sided test is between:

a) 0 and 0.025

b) 0.025 and 0.05

c) 0.05 and 0.1

d) 0.1 and 0.25

e) 0.25 and 1

f) none of the preceding

Q130. A researcher studies a difference between two programming languages. Twelve experts familiar with both languages were asked to write a code for a particular function using both languages and the time for writing those codes was registered. The observations are as follows.

Expert	01	02	03	04	05	06	07	08	09	10	11	12
Lang 1	17	16	21	14	18	24	16	14	21	23	13	18
Lang 2	18	14	19	11	23	21	10	13	19	24	15	29

Construct a 95% C.I. for the mean difference between the first and the second language. Do we have any evidence that one of the languages is preferable to the other (i.e. the average time to write a function is shorter)?

a)$[-1.217, 2.550]$, indication that language 2 is better

b)$[-1.217, 2.550]$, no evidence that any of them is better

c)$[-1.217, 2.550]$, indication that language 1 is better

d)$[-2.86, 4.19]$, no evidence that any of them is better

Q131. For a set of 12 pairs of observations on (x_i, y_i) from an experiment, the following summary for x and y is obtained:

$$\sum_{i=1}^{12} x_i = 25,\ \sum_{i=1}^{12} y_i = 432,\ \sum_{i=1}^{12} x_i^2 = 59,\ \sum_{i=1}^{12} x_i y_i = 880.5,\ \sum_{i=1}^{12} y_i^2 = 15648.$$

The estimated value of y at $x = 5$ from the least squares regression line is:

a)27.78 b)47.77 c)41.87 d)55.97 e)none of the preceding

Q132. Assuming that the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ is appropriate for $n = 14$ observations, the estimated regression line is computed to be

$$\hat{y} = 0.66490 + 0.83075x.$$

Given that $S_{yy} = 4.1289$ and $S_{xy} = 4.49094$, compute the estimated standard error for the slope.

a) 0.3176 b) 0.0783 c) 0.0855 d) 0.0073 e) none of the preceding

Q133. An engineer wants to study the variability of the production process. A sample of size $n = 10$ was taken every 2 hours. After $m = 25$ preliminary samples, one obtains $\overline{\overline{x}} = 30.2$ and $\overline{r} = 7.695$. Determine the lower and the upper control limits for an R chart.

a)LCL$= 1.69$ and UCL$= 13.70$

b)LCL$= 27.81$ and UCL$= 32.59$

c)LCL$= 26.90$ and UCL$= 33.51$

d)LCL$= 28.69$ and UCL$= 31.71$

e)none of the preceding

Q134. Consider the same set-up as in **Q133**. Determine the lower and the upper control limits for an $\overline{X}$ chart from $\overline{R}$.

a)LCL$= 1.69$ and UCL$= 13.70$

b)LCL$= 27.81$ and UCL$= 32.59$

c)LCL$= 26.90$ and UCL$= 33.51$

d)LCL$= 28.69$ and UCL$= 31.71$

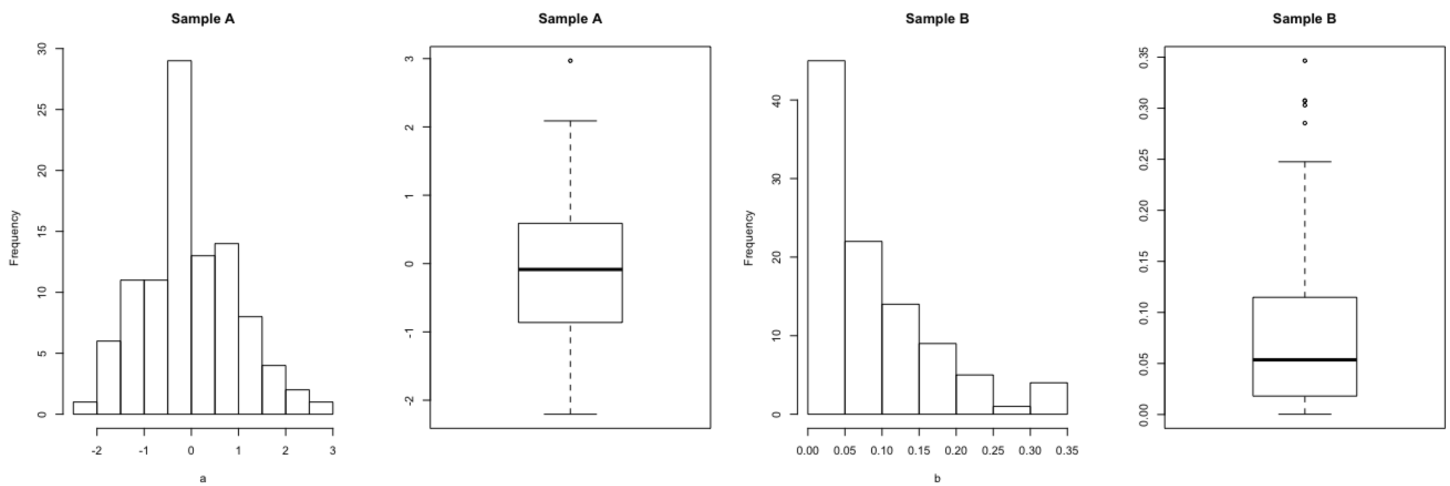
e)none of the preceding

Q135. A company manufactures computers. To control their quality, 50 computers are tested every day. The number of defectives for 15 consecutive days are:

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Xi	3	2	5	1	4	1	6	3	5	0	6	2	4	1	7

What are the lower and upper control limits for the mean number of defectives?

Q136. The following charts show a histogram and a boxplot for two samples, A and B. Based on these charts, we may conclude that



a)only A arises from a normal population

b)only B arises from a normal population

c)both A and B arise from a normal population

Q137. We have a dataset with $n = 25$ pairs of observations (x_i, y_i), and

$$\sum_{i=1}^{n} x_i = 325.000, \ \sum_{i=1}^{n} y_i = 658.972,$$

$$\sum_{i=1}^{n} x_i^2 = 5525.000, \ \sum_{i=1}^{n} x_i y_i = 11153.588, \ \sum_{i=1}^{n} y_i^2 = 22631.377.$$

Note that $t_{0.05/2}(23) = 2.069$. The point estimate for the slope of the regression line is

a)1.99 b)-1.99 c)0.49 d)0.59 e)none of the preceding

Q138. We have a dataset with $n = 25$ pairs of observations (x_i, y_i), and

$$\sum_{i=1}^{n} x_i = 325.000, \ \sum_{i=1}^{n} y_i = 658.972,$$

$$\sum_{i=1}^{n} x_i^2 = 5525.000, \ \sum_{i=1}^{n} x_i y_i = 11153.588, \ \sum_{i=1}^{n} y_i^2 = 22631.377.$$

Note that $t_{0.05/2}(23) = 2.069$. The point estimate for the intercept of the regression line is

a)1.99 b)-1.99 c)0.49 d)0.59 e)none of the preceding

Q139. We have a dataset with $n = 25$ pairs of observations (x_i, y_i), and

$$\sum_{i=1}^{n} x_i = 325.000, \ \sum_{i=1}^{n} y_i = 658.972,$$

$$\sum_{i=1}^{n} x_i^2 = 5525.000, \ \sum_{i=1}^{n} x_i y_i = 11153.588, \ \sum_{i=1}^{n} y_i^2 = 22631.377.$$

Note that $t_{0.05/2}(23) = 2.069$. What is the prediction of y for $x = 30$?

a)60.19 b)16.67 c)30 d)30.54 e)none of the preceding

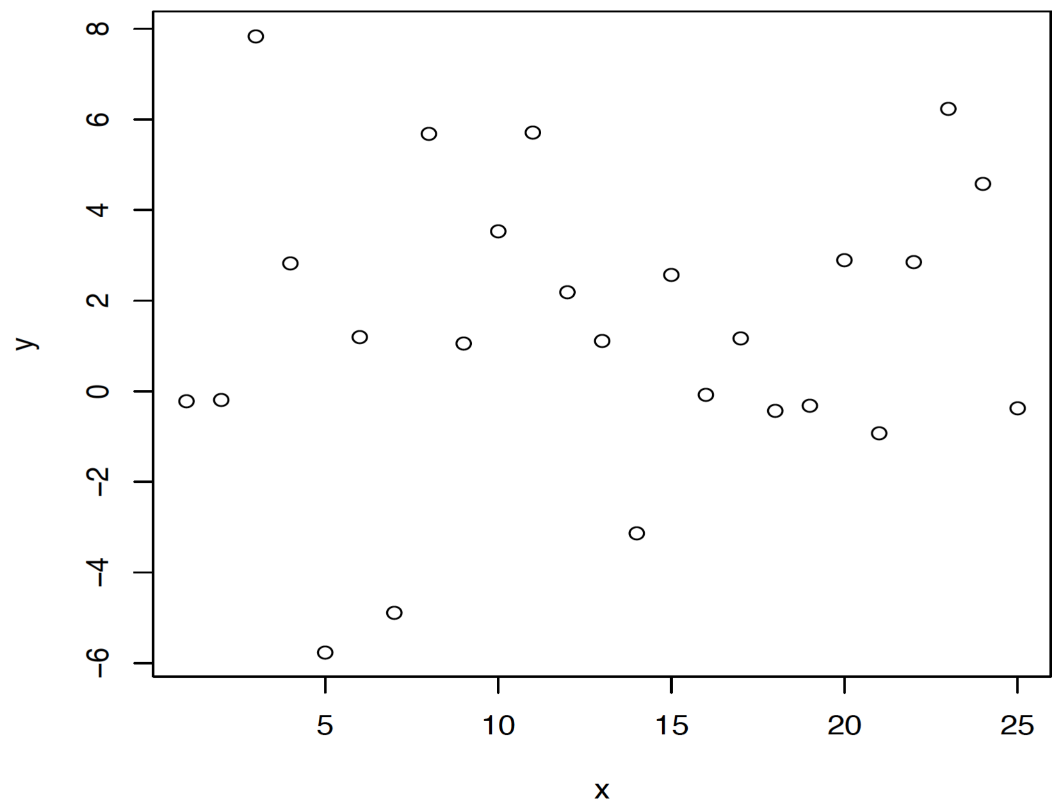
Q140. We have a dataset with $n = 25$ pairs of observations (x_i, y_i), and

$$\sum_{i=1}^{n} x_i = 325.000, \ \sum_{i=1}^{n} y_i = 658.972,$$
$$\sum_{i=1}^{n} x_i^2 = 5525.000, \ \sum_{i=1}^{n} x_i y_i = 11153.588, \ \sum_{i=1}^{n} y_i^2 = 22631.377.$$

Note that $t_{0.05/2}(23) = 2.069$. Is the linear regression significant?

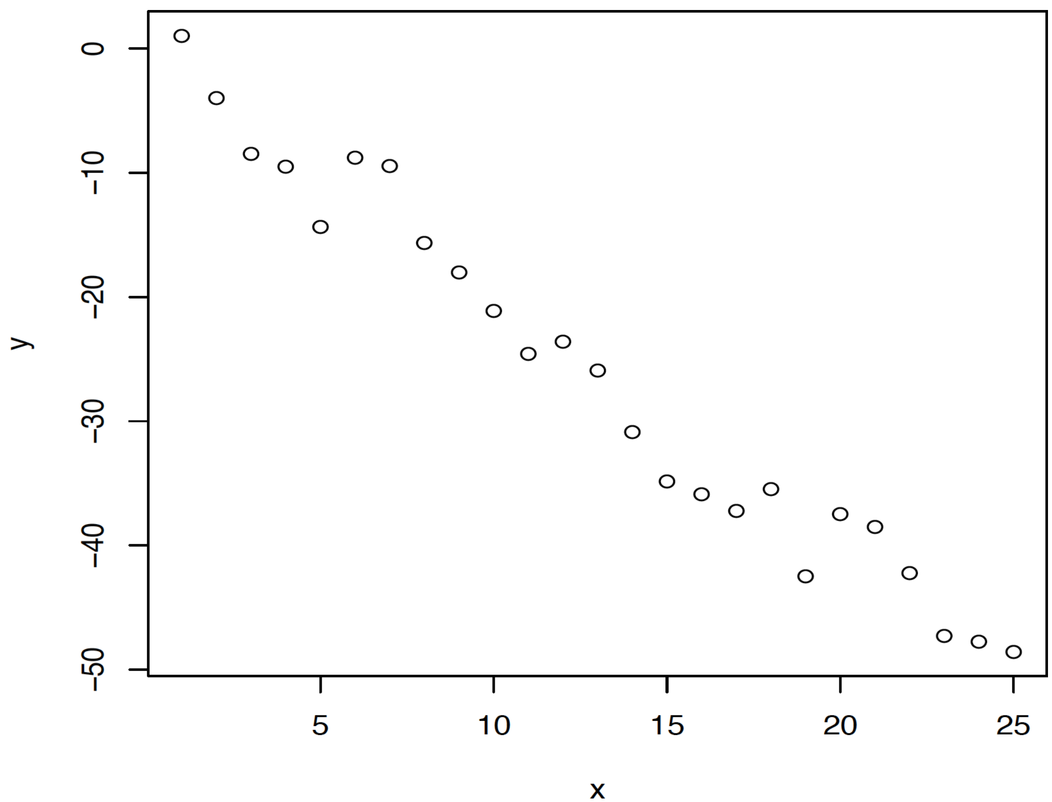
Q141. For the following data the correlation coefficient is most likely to be

a)0.01 b)0.98 c)-0.5 d)-0.98



Q142. For the following data the correlation coefficient is most likely to be

a)0.01 b)0.98 c)-0.5 d)-0.98



Q143. A company employs 10 part-time drivers for its fleet of trucks. Its manager wants to find a relationship between number of km driven (X) and number of working days (Y) in a typical week. The drivers are hired to drive half-day shifts, so that 3.5 stands for 7 half-day shifts.

The manager wants to use the linear regression model $Y = \beta_0 + \beta_1 x + \epsilon$ on the following data:

	1	2	3	4	5	6	7	8	9	10
x	825	215	1070	550	480	920	1350	325	670	1215
y	3.5	1.0	4.0	2.0	1.0	3.0	4.5	1.5	3.0	5.0

Note that $\sum x_i^2 = 7104300$, $\sum y_i^2 = 99.75$, and $\sum x_i y_i = 26370$. What is the fitted regression line?

Q144. Using the data from question **Q143**, what value is the correlation coefficient of x and y closest to?

a) 0.437 b) 0.949 c) 0.113 d) 1.123 e) none of the preceding

Q145. We want to test significance of regression, i.e. $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$. The value of the appropriate statistic and the decision for $\alpha = 0.05$ is:

a) 8.55; do not reject H_0

b) 2.31; reject H_0

c) 8.55; reject H_0

d) 2.31; do not reject H_0

e) none of the preceding

Q146. 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 $n = 20$,

$$\sum y_i = 12.75,\ \sum y_i^2 = 8.86,\ \sum x_i = 1478\ \sum x_i^2 = 143,215.8\ \sum x_i y_i = 1083.67.$$

a) Calculate the least squares estimates of the slope and intercept. Estimate σ^2.

b) Use the equation of the fitted line to predict what pavement deflection would be observed when the surface temperature is 90F.

c) Give a point estimate of the mean pavement deflection when the surface is 85F.

d) What change in mean pavement deflection would be expected for a 1F change in surface temperature?

Q147. Consider the data from **Q146**.

a) Test for significance of regression using $\alpha = 0.05$. Find the p-value for this test. What conclusion can you draw?

b) Estimate the standard errors of the slope and intercept.

Q148. Solve this question using `R`.

a) Generate a sample `x` of size $n = 100$ from a normal distribution;

b) Define `y=1+2*x+rnorm(100)`;

c) Plot scatter plot;

d) Find the estimators of the regression parameters and add the line to the scatter plot;

f) Compute the correlation coefficient

g) Plot the residuals;

h) Comment on your results.

Q149. We have $m = 5$ preliminary samples of size $n = 3$ (some numbers have unfortunately been erased by accident by a clumsy co-op student):

i	$x_{i,1}$	$x_{i,2}$	$x_{i,3}$	$\bar{x}_i$	r_i	s_i
1	27.1	29.4		27.9		1.3
2	30.6	32.5	32.4	31.83	1.9	1.07
3	25.7	35.5	30	30.4		4.91
4	31.1	23.2	25	26.43	7.9	
5	24.1	34.2	27.4	28.57	10.1	5.15
			total:	145.13	32	16.57

What is the control chart (give the interval) for $\overline{X}$ from R?

Q150. We have $m = 5$ preliminary samples of size $n = 3$ (some numbers have unfortunately been erased by accident by a clumsy co-op student):

i	$x_{i,1}$	$x_{i,2}$	$x_{i,3}$	$\bar{x}_i$	r_i	s_i
1	27.1	29.4		27.9		1.3
2	30.6	32.5	32.4	31.83	1.9	1.07
3	25.7	35.5	30	30.4		4.91
4	31.1	23.2	25	26.43	7.9	
5	24.1	34.2	27.4	28.57	10.1	5.15
			total:	145.13	32	16.57

What is the control chart (give the interval) for $\overline{X}$ from S?