You can only use the faculty approved calculator with sticker.
Pens, pencils, erasers, and straight edges only. No crib sheets. NO CELL PHONES.
If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Justify your responses.

1. A security camera is used to monitor a particular surveillance area to report whether there is any threat present or not. The probability that the camera correctly detects an actual threat as "threat found" is 0.8. The probability that the camera correctly reports no threat as "no threat found" is 0.9. Suppose the probability of an actual threat being present in the surveillance area is 0.6.

   a. What is the probability that the camera reports "threat found"? (4 Marks)
   b. Given that the camera reports "threat found", what is the probability that there are actual threats in the surveillance area? (3 Marks)

Answer parts c) and d) by assuming that an actual threat is always present (i.e., probability of threat presence is 1)

   c. What is the probability that the camera reports "threat found"? (1 Mark)
   d. An identical second camera is used to increase the accuracy of threat reporting capability. The cameras are operating independently and monitoring the same surveillance area. Find the probability that the combined system report the threat accurately (at least one camera should report "threat found"). (2 Marks)

2. Continuous Random variables.
   a. Assume a random variable X is normally distributed with a mean of 10 and a variance of 16. Find A such that $P(X>A)=0.7$. (4 marks)

   Now consider the probability function g(x)

$$g(x)=\begin{cases} kx & 0<x<10 \\ 0 & elsewhere \end{cases}$$

   b. Find the value of k. (3 marks)
   c. Calculate $P(5<X<10)$. (3 marks)

3. Parts come off the line one by one. Since the line has just been started up it is not very good and the probability that any one part is defective is 0.1. Assume that the defectiveness of different parts is independent of one another.
   a. What is the probability that part 14 or 15 is the 1st defective part?(3 marks)
   b. What is the mean number of parts that one will examine before the first defective part? (2 marks)
   c. What is the probability that part 14 or 15 will be the second defective part? (3 marks)
   d. Consider now a new random variable Z with the following pmf f(z)

z	-1	0	1	2
f(z)	0.1	0.4	0.2	0.3

Calculate the mean of Z, E[Z]. (2 marks)

You can only use the faculty approved calculator with sticker.
Pens, pencils, erasers, and straight edges only. No crib sheets. NO CELL PHONES.
If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Justify your responses.
Good Luck

1. A test to detect pollutants is 90% accurate when pollutants are present in a sample and 96% accurate when pollutants are not present. This test is used on a set of samples of which 65% are polluted.
   a. For a particular sample, what is the probability that the test says pollutants are present? **(5 marks)**
   b. If the test says that pollutants are not present in a particular sample, what is the probability that in fact there are pollutants in the sample in question? **(5 marks)**

2. Cars moving along a highway come to an intersection where they can choose to go in one of two directions, A or B. They will chose direction A with probability 0.2, and direction B the other times. Assume that the behavior of different cars is independent.
   a. What is the probability that the 4^{th} car along the highway is the first to choose direction B? **(3 marks)**
   b. Of the next 10 cars what is the probability that more than 3 choose direction A? **(3 marks)**
   c. Direction B leads to a 4 car ferry. The ferry will wait until five cars are on the ferry and then it will leave. If now there are no cars on the ferry determine the probability that the 20^{th} car along the highway will be the one to fill up the ferry. **(4 marks)**

3. Assume X is normally distributed with a mean of 12 and a variance of 16.
   a. Calculate $P(9<X<17)$. **(3 marks)**
   b. Find x_1 such that $P(X>x_1) = 0.6$. **(3 marks)**
   c. Consider an interval for X, $(a<X<b)$. Calculate a and b such that the interval is shortest and $P(a<X<b) = 0.3$. **(4 marks)**

4.

You can only use the faculty approved calculator with sticker.
Pens, pencils, erasers, and straight edges only. No crib sheets. NO CELL PHONES.
If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Justify your responses.
Good Luck

1. In a certain school there are only Electrical Engineering students, Civil Engineering Students and Mechanical Engineering students. These students must take a class where they give a presentation. Students feel that the order of presentations is very important.
    a. If there are 9 students in the class and all give their presentation on one day how many different ways can you arrange the order of the presentations? **(1 mark)**
    b. If there are 18 students in the class but only 9 can give their presentation on Day 1, how many different orders of presentation can you have for Day 1. **(2 marks)**
    c. If the class contains 7 Mechanicals, 6 Civils and 5 Electricals and we require that the 9 individuals who give their presentations on Day 1 must include 3 Mechanicals, 3 Civils and 3 Electricals, how many different orders of presentation are there for Day 1? **(4 marks)**
    d. Suppose that all of the presentations that you counted in c) part are equally likely. Suppose that you are one of the 5 Electrical Students. What is the probability that you will be the first one to give a presentation on Day 1? **(3 marks)**
2. Your factory receives 3 lots of screws. There are 100 screws in Lot 1, 200 in Lot 2 and 300 in Lot 3. Lot 1 contains 1% defective screws. Lot 2 contains 2% defective screws and Lot 3 contains 3% defective screws. All the screws are dumped in one bin and then a screw is selected randomly.
    a. What is the probability that the selected screw came from Lot 1? **(3 marks)**
    b. What is the probability that the selected screw was defective? **(3 marks)**
    c. If the screw was defective what is the probability that it came from Lot 2? **(4 marks)**
3. You work as a lobster fisherman. You set several traps every day. There is a probability of 0.08 that any one of your traps will catch a lobster on a given day. Assume that whether or not a trap catches a lobster is independent of what happens with the other traps.
    a. If you set 10 traps, what is the probability that you will catch more than 2 lobsters? **(3 marks)**
    b. If you catch more than 2 lobsters your boss will let you eat one for dinner. (Note that even if you catch 5 lobsters in one day you still only get one for dinner.)What is the probability that you will be able to eat a lobster in the next 3 days? **(4 marks)**
    c. What is the expected number of days you will have to wait until you will have had 2 lobster dinners? **(3 marks)**

5.

1. Statistical data gathered for Toyota cars that are returned for mechanical failure within their first year of usage show that Engine, Sensors and Breaks are the three main sources for the mechanical failures. Toyota's data provides the following information for:

   E = {Engine failures}
   B = {Break failures}
   S = {Sensor problems}

   $P(E) = 0.35$

   $P(E \cap B') = 0.29$

   $P(B \cap S) = 0.09$

   $P(B' \cap S') = 0.25$

   $P((E \cap S)/B) = 0.2$

   $P(S/E) = 9/35$

   $P(B') = 0.75$

   $P(E/B) = 6/25$

   For the given information answer the following questions **(1 mark each)**

   a) $P(E \cap S) = ?$

   b) $P(B) = ?$

   c) $P(E \cap S \cap B') = ?$

   d) $P(E \cap S \cap B) = ?$

   e) $P(B \cap E') = ?$

(HINT: A Venn diagram may be useful here)

2. Probability distribution for the *cm* of snow fall in Montreal in a typical light snow fall, X, is given by:

$$f(x) = \begin{cases} \dfrac{2x-2}{k} & 1 \le x \le 4 \\ \dfrac{20-2x}{54} & 4 \le x \le 10 \end{cases}$$

   If the snow removal cost to city is given by the random variable Y defined as $Y = 5000x - 2000$, find:

   a) the value of k **(2 marks)**
   b) the expected value and the variance of the random variable Y **(3 marks)**

3. A very large batch of components has arrived at a distributor. The batch can be thought of as "acceptable" only if the portion of defective components is 10% or less. The distributor decides to randomly select 10 components and test them. The distributor will label the entire batch as "acceptable" if at most two parts are defective.
   a. If the actual portion of defectives in a batch is 5% calculate the probability of incorrectly labeling that batch as not acceptable. **(1 marks)**
   b. If the actual portion of defectives in a batch is 15% calculate the probability of incorrectly labeling the batch as acceptable. **(1 marks)**
   c. If the actual portion of defectives in a batch is 15%, what is the expected number of defectives in the sample of 10? **(1 marks)**
   d. The distributor is concerned that errors in labeling of batches is costing too much, in particular the labeling of batches as acceptable when they have more than 10% defectives is very costly. Make concrete suggestions on how things might be improved. Be specific. **(2 marks)**

4. A certain store has an express cashier and a super-express cashier. Let X_1 be the number of customers in the express cashier line. Let X_2 be the number of cashiers in the super-express cashier line. Suppose that the joint pmf of X_1 and X_2 is

X_2 \ X_1	0	1	2	3
0	0.08	0.07	0.04	0.00
1	0.06	0.15	0.05	0.04
2	0.05	0.04	0.10	0.06
3	0.00	0.03	0.04	0.07
4	0.00	0.01	0.05	0.06

Thus X_1 takes on values 0,1,2 and 3 and X_2 takes on values 0,1,2,3 and 4.

   a. Find the probability that there is exactly one person in each line. **(1 mark)**
   b. Let A be the event that there are at least two more customers in one of the lines than the other. Find the probability of event A. **(2 marks)**
   c. What is the probability that the total number of customers in both lines is 4. **(2 marks)**
   d. If we know that there are two people in the express cashier line what is the probability that there are at least 3 people in the super express cashier line? **(2 marks)**
   e. Are X_1 and X_2 independent? Justify your response. **(3 marks)**

Total 25 marks

Midterm – ENGR 371 – July 1999

Pens, pencils, erasers, and straight edges only. Calculator. No crib sheets.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Justify your responses.

1. It may seem counterintuitive, but the distribution of the first significant digit in many collections of numbers is NOT uniform on the integers 1 through 9, but follows Benford's distribution which is as follows:

$$f(x) = \begin{cases} \log_{10}(1+\frac{1}{x}) & x = 1, 2, ...9 \\ 0 & \text{otherwise} \end{cases}$$

One use of this distribution is in auditing financial records. The idea is that if the books have been artificially altered, the distribution of the first significant digit will differ markedly from what is predicted by Benfords's distribution. Suppose that Revenue Canada is auditing the records of a large company.

(a) If the numbers in the company's records follow Benford's distribution, what is the probability that a randomly chosen number from the company's records has a first significant digit of 1? (2 marks)

(b) If 1000 numbers are selected randomly from the company's records and if Y is the number of those 1000 that start with 1, what is the distribution of Y? Be sure to state all parameters of the distribution. (2 marks)

(c) Of the 1000 numbers above what is the expected number of those that start with 1? (2 marks)

(d) If of the 1000 numbers sampled from the financial data sheets 120 started with 1 would you conclude that there is evidence that the financial records have been artificially altered? Explain. (2 marks)

(e) If you were to randomly select numbers from the financial records one at a time, and Z was the number you needed to select until you had 5 numbers starting with 1, what distribution would Z have? Be sure to state clearly all parameters of Z's distribution. (2 marks)

2. Shoppers are questioned as to the acceptability of three brands of soap: A, B and C. Of those questioned 67 found brand A acceptable, 43 found brand B acceptable, and 38 found brand C acceptable. Furthermore, 21 found both A and B acceptable, 14 found both B and C acceptable and 14 found both A and C acceptable. There were 4 shoppers who found all three brands acceptable. If a total of 108 shoppers were questioned, how many found all three brands unacceptable? (5 marks)

3. The joint density function of two random variables X and Y is

$$f(x,y) = \begin{cases} (x+y)/8 & 0 < x < 2, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Are X and Y independent? Justify your response. (5 marks)

(b) Evaluate $P(|X - Y| > 1)$. (3 marks)

(c) Calculate the expected value of $3X$. (2 marks)

(Exam Total 25 marks)

Midterm – ENGR 371 – February 1999

Pens, pencils, erasers, and straight edges only. No crib sheets.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Justify your responses.

1. Consider a box containing N balls. It contains red balls and blue balls. We will be sampling balls from the box one at a time. As each ball is taken out, it is thrown away and replaced with a blue ball, no matter what the colour of the sampled ball was.

   (a) If to begin with k of the balls are red and the rest blue let X be the number of balls drawn until a red ball is drawn.

   i. Give the distribution of X. (3 marks)

   ii. Give the expected value of X. (2 marks)

   (b) Say the box contains only red balls to begin with. The sampling procedure outlined above is performed. Let S be the total number balls sampled until the box contains only blue balls. Find the expected value of S. (4 marks)

2. The joint density function of two random variables X and Y is

$$f(x,y) = \begin{cases} 6x & 0 < x < 1, 0 < y < 1 - x \\ 0 & \text{otherwise} \end{cases}$$

   (a) Are X and Y independent? Justify your response. (2 marks)

   (b) What is the probability that X is bigger than Y by at least 0.25? (4 marks)

   (c) Find $P(X < 0.2 | Y = 0.3)$. (4 marks)

3. A chemical company is interested in the impurity levels of the chemicals it manufactures. Previous experience indicates that one in one hundred of its chemical batches has an impurity level which is too high. A new state of the art laser based device for measuring impurities is purchased, but it is not foolproof. It will give a reading of "too high" for 5% of the batches which actually have a satisfactory impurity level. This is called a "false positive" or a "false alarm". On the other hand, it will falsely indicate satisfactory impurity for 2% of the batches that actually have too high an impurity level. This is called a "false negative" or a "miss".

   (a) If the test says the impurity level is too high, what is the probability that it really is too high? (4 marks)

   (b) If the test says the impurity level is satisfactory, what is the probability that it really is satisfactory? (2 marks)

(Exam Total 25 marks)

Some Useful Equations

Midterm – ENGR 371 – February 1997

Pens, pencils, erasers, and straight edges only. No calculators. No crib sheets.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Justify your responses.

1. An electronic assembly consists of two subsystems say A and B. From previous testing procedures, the following probabilities are assumed to be known:
   P(A fails) $= 0.2$
   P(B fails and A doesn't fail) $= 0.15$
   P(both A and B fail) $= 0.15$
   Evaluate the following:

   (a) P(A fails|B has failed).

   (b) P(A fails and B doesn't fail).

   (Total 7 marks)

2. Let X denote the thrust and Y denote the mixture ratio, which are two major characteristics of rocket engine's performance. Suppose that X and Y are continuous random variables with joint probability distribution function:

$$f(x,y) = \begin{cases} k(x+y-2xy) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Find k.

   (b) Determine whether X and Y are independent.

   (c) Compute $P(X < \frac{1}{2} | Y = \frac{1}{2})$.

   (Total 9 marks)

3. The total time, measured in units of 100 hours that an Arts and Science student runs his/her stereo over a period of one year is a continuous R.V. X with pdf

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

   (a) Determine EX

   (b) Determine the variance of X.

   (c) Define $Y = 60X^2 + 30X$. Determine the mean of Y.

   (d) Determine the variance of Y.

   (Total 9 marks)

   (Exam Total 25 marks)

Midterm – ENGR 371 – February 1996

Pens, pencils, erasers, and straight edges, calculators allowed. No crib sheets.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Show all your work and justify all your answers. Marks are given for how an answer is arrived at not just the answer itself.

1. A town is protected from floods by a reservoir dam designed for a 50-year flood; that is, the probability that the reservoir will overflow in a year is 1/50 or 0.02. The town and reservoir are located in an active seismic region; annually the probability of occurance of a destructive earthquake is 5 %. During such an earthquake, it is 20 % probable that the dam will be damaged, thus causing the reservoir water to flood the town. Assume that the occurances of natural floods and earthquake are statistically independent.

   (a) What is the probability of an earthquake-induced flood in a year?

   (b) What is the probability that the town is free from flooding in any one year?

   (c) If the occurence of an earthquake is assumed in a given year, what is the probability that the town will be flooded that year?

   (7 marks)

2. A random variable Y takes on the values 1, 2 or 3 with probabilities $(1+3k)/3$, $(1-2k)/3$ and $(1-k)/3$, respectively.

   (a) For what values of k is this a legitimate probability function?

   (b) Find the cumulative distribution function.

   (c) Find the mean and variance of Y.

   (7 marks)

3. In an atomic reaction, N particles are fired at a surface. Each particle releases a random number X_i of particles. The random variables $X_1, X_2, ..., X_N$ are independent, each having a mean μ and variance σ^2. The total number of particles released is: $X = X_1 + X_2 + X_3 + ... + X_N$

   (a) Assume that it is known that $N = 50$. Find the mean and variance of X.

   (b) Now assume that all that is known about N is that it is a random variable with mean m and variance s^2. N is independent of the X_i. Find the mean and variance of X.

   (4 marks)

4. Consider the following two discrete random variables, X and Y, with joint distribution given in the following Table.

f(x,y) X	Y			
	y=-1	y=0	y=1	y=3
x=-1	1/24	1/12	1/12	1/24
x=0	1/12	1/6	1/6	1/12
x=1	1/24	1/12	1/12	1/24

(a) Determine the marginal distribution of X.

(b) Determine the marginal distribution of Y.

(c) Are X and Y independent?

(d) Determine the conditional distribution $f(Y|X=0)$.

(e) Determine $P(X > Y)$.

(7 marks)

Some Useful Equations

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$$

$$\sum_{i=0}^{n} r^i = \frac{1-r^{n+1}}{1-r}$$

$$\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$$

$$f(x) = \binom{n}{x} p^x q^{n-x}$$

$$f(x) = \binom{x-1}{k-1} p^k q^{x-k}$$

$$f(x) = pq^{x-1}$$

$$f(x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$\int_0^{\infty} x^{\alpha-1} e^{-x} dx = (\alpha-1)!$$

$$f(x) = \frac{1}{\beta} e^{\frac{-x}{\beta}}$$

Midterm – ENGR 371 – November 1994

Pens, pencils, erasers, and straight edges only. No calculators. No crib sheets.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Show all your work and justify all your answers. Marks are given for how an answer is arrived at not just the answer itself.

1. You are given two independent random variables X and Y with the following marginal distributions:

$$g(x) = \begin{cases} 0.5e^{-0.5x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$h(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Two new random variables are defined:

$$U = 3X + Y$$

$$V = X + 2Y$$

(a) What is the expected value of X? (2 marks)

(b) What is the expected value of Y? (2 marks)

(c) What is $E(XY)$? (2 marks)

(d) Find the joint probability distribution of X and Y. (2 marks)

(e) Find the joint probability distribution of U and V. (9 marks)

(f) What is the probability of $(U + 2V) < 5$? (6 marks)

(g) What is the expected value of U? (2 marks)

2. A travelling salesperson based in Montréal must visit Hull, Québec City, Trois Rivières, Chicoutimi and Sherbrooke, over the next two weeks in one long trip. In making travel plans he/she would like to select the route that minimizes travel time. To this end he/she figures out the total distance required for each of the possible routes.

(a) How many possible routes are there? (5 marks)

(b) If Chicoutimi and Sherbrooke must be visited one after the other (but not neccessarily in that order), how many routes would there be? (5 marks)

3. If the probability is 0.1 that any one person will dislike the taste of a new toothepaste, what is the probability that exactly five of eighteen selected persons will dislike it? (5 marks)

(Total 40 marks)

Some Useful Equations

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$$

$$\sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1 - r}$$

$$\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$$

$$f(x) = \binom{n}{x} p^x q^{n-x}$$

$$f(x) = \binom{x-1}{k-1} p^k q^{x-k}$$

$$f(x) = pq^{x-1}$$

$$f(x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$\int_0^{\infty} x^{\alpha-1} e^{-x} dx = (\alpha - 1)!$$

$$f(x) = \frac{1}{\beta} e^{\frac{-x}{\beta}}$$

QUESTIONS

1. The life in hours of a battery is known to be normally distributed, with standard deviation of 1.5 hours. A random sample of 12 batteries has a sample mean life time of 50 hours.
   a. Test the hypothesis that the mean battery life is 50.5 hours? (Use $\alpha = 0.05$) (5 marks)
   b. What is the P-value for the test in part (a)? (5 marks)
   c. Find the β-error for the test in part (a), if the true mean life is 52 hours? (5 marks)
   d. What sample size is required to ensure that β does not exceed 0.10, if the true mean life if 52 hours? (5 marks)

2. A random sample of 36 observations has been drawn from a normal distribution with mean of 50 and standard deviation of 12.
   a. Find the probability that the sample mean is in the interval of $47 \le \bar{X} \le 53$. (7 marks)
   b. Is the assumption of normality important? Why? (3 marks)

3. The tar content in 9 samples of cigar tobacco follows 5.40, 5.67, 5.79, 6.85, 6.92, 5.70, 6.08, 5.48, and 5.44. Assume the tar content has a normal distribution.
   a. Calculate the sample mean and variance (4 marks)
   b. Estimate 95% confidence upper bound of the population mean (6 marks)

4. For a certain virus, if a person is exposed to the virus there is a 10% chance that the person will be infected. Assume that persons who were infected are independent of one another.
   a. If we examine exposed people one by one, find the probability that we must examine 5 or more people before we find two who are infected. (5 marks)
   b. What is the mean number or the expected number of people we must examine before we find two people who are infected? (5 marks)

5. A hydraulic system contains two pumps and two valves. Let P_i be the event that pump i works (here i=1,2). Similarly let V_j be the even that valve j works (here j=1,2). The entire system will work if at least one pump and at least one valve work. Let F be the event that the whole system works. Write down F in terms of the P_i and V_j. (5 marks)

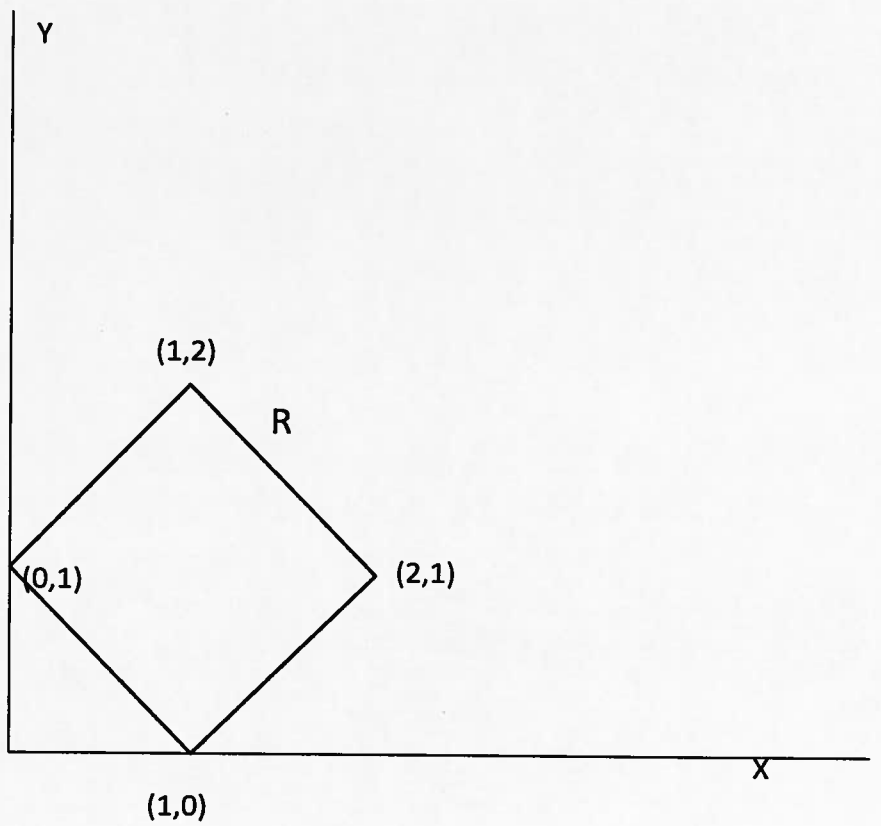
6. Four friends are given tickets to a football game. The tickets give seats all in the same row of 12 seats. However the tickets are for random seats in that row. What is the probability that the four friends will be seated in consecutive seats (i.e. all four seated one beside the other)? (5 marks)

7. X and Y have the following joint pdf.

$$f(x,y) = \begin{cases} kx & (x,y) \in R \\ 0 & otherwise \end{cases}$$

Here R is the interior of the diamond region shown below.

a) Determine the constant k. (5 marks)
b) Determine P(X|Y=1.5). Be sure to state the limits. (5 marks)
c) Find the marginal distribution of Y. Be sure to state the limits. (5 marks)
d) Are X and Y independent? Justify your response. (5 marks)



8. The waiting time in a bank is exponentially distributed with a mean of 10 minutes.
   a. What is the probability that a customer will spend more than 15 minutes in the bank? (5 Marks)
   b. What is the probability that a customer will spend more than 15 minutes in the bank given that he/she spends 10 minutes already in the bank? (5 Marks)

QUESTIONS

1. The world health organization wants to investigate the weights of children. A random sample of size 16 is collected and the weights are 20.5, 20, 20.8, 21, 19, 18, 23, 21.5, 22.5, 18.5, 19, 19.5, 25, 23, 21, and 24 kg. Assume that the population distribution is normal.
   (a) Test the hypothesis on the mean: H_0: μ=20 kg and H_1: μ>20 kg. Use α=0.05. **(10 marks)**
   (b) Test the hypothesis on the variance: H_0: σ^2=2 kg^2, and H_1: $\sigma^2 \neq 2$ kg^2. Use α=0.005. **(10 marks)**

2. Final scores for ENGR 371 students follow a normal distribution with a standard deviation of 10. A random sample of 20 students is selected and it is found that the sample mean of the final scores is 65. Find the 95% confidence interval for the final scores of the population mean of ENGR371 grades. **(10 marks)**

3. The joint probability function of random variables X and Y is given by
   $f(x,y) = ke^{-2x-3y}$, 0<x<y.
   (a) Find k **(5 marks)**
   (b) Find P(1<X<2) **(5 marks)**
   (c) Find the mean of $Z = 2X + 3Y + 1$ **(5 marks)**
   (d) Are the random variables X and Y independent? Justify your answer **(5 marks)**

4. A product is made by three machines, and machine #1, #2 and #3 produce 50%, 25% and 25% of the product, respectively. Based on the previous knowledge, it is known that 2%, 1% and 3% of the products made by machine #1, #2 and #3, respectively are defective.
   (a) If a finished product is randomly selected, what is the probability that the selected product is defective? **(10 marks)**
   (b) If a finished product is randomly selected and is found to be defective, which machine is most likely to produce this defective product? **(10 marks)**

5. In hockey games, Concordia has a probability of 0.55 of beating McGill in each game that they play. Assume that the outcomes of games are independent of one another
   (a) Determine the probability that it takes Concordia 11 games until it has beaten McGill 6 times. **(5 marks)**
   (b) McGill and Concordia both reach the championships. There are 7 games and the team that wins 4 or more wins. What is the probability that Concordia wins the championship? Hint: be careful. **(5 marks)**

6. A random sample of size n_1=16 is selected from a normal population with a population mean of 75 and a population standard deviation of 8. A second random sample of size n_2=9 is taken from another normal population with population mean of 70 and population standard deviation of 12. Assume that individual samples from the different populations are independent of each other. Let $\bar{X}_1$ and $\bar{X}_2$ be the two sample means. Find the probability of $3.5 \leq \bar{X}_1 - \bar{X}_2 \leq 5.5$. **(5 marks)**

7. We measure the resistance of 2500 resistors that are each nominally 100 ohms. These come from a very large lot of resistors. Assume that these 2500 resistors are selected so that they are a random sample. We find that the sample mean is 99.97 ohms and the sample standard deviation is 1.1 ohms. Calculate a 95% confidence interval on the population mean of 100 ohm resistors from this lot. **(10 marks)**

8. Vehicles enter a certain on ramp to the highway according to a Poisson Process with λ=17 vehicles/hour. Determine the probability that there are less than 4 minutes between two consecutive vehicles. **(5 marks)**

Concordia University
Department of Electrical and Computer Engineering
ENGR 371 - Probability and Statistics

Drs. Lynch and Ghrayeb

Final Examination — April 29, 2010

1) **(10 Marks)** In a particular area, there are just three makes of cars available: Ford (60% of the total), Renault (30%) and Honda (10%). The carburetors on Fords have a 0.09 probability of failure, those on Renaults have a 0.05 chance and those on Hondas have a 0.02 chance of failure. If a car with a failed carburetor is selected at random, what is the probability that it is a Ford?

2) **(10 Marks)** A multiple-choice quiz has 200 questions each with 4 possible answers of which only 1 is the correct answer. What is the probability that sheer guesswork yields from 25 to 30 correct answers for 80 of the 200 problems about which the student has no knowledge? **[HINT: approximation may be helpful here]**

3) **(10 Marks)** Let X and Y represent two random variables whose joint density function is

$$f_{XY}(x,y) = \begin{cases} 24xy, & 0 \le x \le 1, 0 \le y \le 1, \quad x+y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

   a) Find $f_Y(y)$.
   b) Are X and Y independent? Explain.
   c) Find $P\left(Y < \frac{1}{8} | X = \frac{3}{4}\right)$.

4) **(10 Marks)** Two independent experiments are being run in which two different types of paints are compared. Eighteen specimens are painted using type A and the drying time, in hours, is recorded on each. The same is done with type B.
   a) Suppose that the population standard deviations are both known to be 1.0. Assuming that the mean drying time is equal for the two types of paint, find $P(\bar{X}_A - \bar{X}_B > 1.0)$, where $\bar{X}_A$ and $\bar{X}_B$ are average drying times for samples of size $n_A = n_B = 18$.
   b) Repeat Part (a) when the population standard deviations are unknown, but they are still equal and estimated to be $s = 0.96$.

5) **(10 Marks)** A random sample of 100 automobile owners shows that, in the province of Quebec, an automobile is driven on the average 23,500 kilometers per year with a standard deviation of 3900 kilometers.
   a) Construct a 99% confidence interval for the average number of kilometers an automobile is driven annually in Quebec.
   b) What can we assert with 99% confidence about the possible size of our error if we estimate the average number of kilometers driven by car in Quebec to be 23,500 kilometers per year?

6) **(10 Marks)** The life in hours for a battery is known to be normally distributed with standard deviation $\sigma = 1$ hours. A random sample of 16 batteries has a mean life of $\bar{x} = 20.3$ hours.
   a) Test $H_0 : \mu = 20$ versus $H_1 : \mu > 20$ for $\alpha = 0.05$.
   b) What is the P-value of the test in part (a)?
   c) What is the β error for the test in part (a) if the true mean is 20.75 hours?

8

QUESTIONS

1. A small printing company claims that a certain refilled laser printer cartridge can print minimum of 5000 pages. In order to verify the company's claim, Concordia University conducted a study among 50 refilled printer cartridges purchased form this company and found that cartridges can print average 4870 pages with a population standard deviation of 90 pages. Based on the given information
   a. Compute a 95% confidence interval for the population mean. **(6 points)**
   b. In order to reduce the error in estimating (half-width of) the population mean (e) by one half, how many more sample should be tested? **(4 points)**

2. During a Rock concert, failure of an amplifier is not an option. For the concert total 5 similar kinds of amplifiers are used. Based on the previous observations, it is known that an amplifier can perform without a failure with an average of 2 hours. Failures occasionally happen during the concert and the time that an amplifier works before it fails is found to be exponentially distributed. If the organizer likes to achieve 95 percent reliability during the upcoming concert which last about 2 hours, how many amplifier they should carry during the concert? **(10 points)**

3. Let Y denotes the diameter of a metal-pipe and X is the diameter of a plastic inserted in the joint to make a weatherproof system in an airplane. After scaling both X and Y, we obtain the following joint probability distribution for the system

$$f(x,y)=\begin{cases}\frac{1}{y} & 0<x<y<1\\ 0 & else\end{cases}$$

   a. If the designer wants the scaled tolerance (Y - X) between diameters not to be less than 0.05 and more than 0.10, find the probability of the system to comply with the given tolerance limits. In another word find
      $P(0.05 < Y-X < 0.10)$ **(7 points)**
   b. If we select 2 systems randomly, what is the probability that none of them violates the given tolerance specifications? **(3 points)**

4. From a group of 5 mechanical engineering (ME) and 3 industrial engineering (IE) students, how many committees of size 4 are possible?
   a. With no restrictions **(3 points)**
   b. Committee should include 2 ME and 2 IE **(2 points)**

5. Consider continuous random variables X and Y with the following joint probability density function:

$$f(x,y)=\begin{cases}\frac{1}{2} & x<y<x+2;\ -x<y<-x+2\\ 0 & else\end{cases}$$

   a. For

$$u = \frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}}$$
$$v = \frac{-x}{\sqrt{2}} + \frac{y}{\sqrt{2}}$$

Determine the joint distribution of U and V. Call it g(u,v). **(7 marks)**

b. Determine region of support of g(u,v). That is determine the region in the (u,v) plane where g(u,v) is not zero. **(3 marks)**.

6. We wish to study the flexural strength of a certain type of concrete beam. A random sample of size 30 is taken with the following strengths in Mega Pascals (MPa):

5.9	7.2	7.3	6.3	8.1	6.8	7.0	7.6	6.8	7.1
6.5	7.0	6.3	7.9	9.0	8.2	8.7	7.8	9.7	8.1
7.4	7.7	9.7	7.8	7.7	11.6	11.8	11.3	10.7	9.1

Assume that the population distribution is Gaussian.

a. Calculate the Sample Mean. **(2 marks)**
b. Calculate the Sample Median. **(1 mark)**
c. Calculate the Sample Standard Deviation. **(2 marks)**

If a beam is stronger than 8 MPa it has passed its strength test and can be considered a success.

d. Calculate the probability that the population mean is greater than 8 MPa. **(3 marks)**
e. Provide an estimate of what portion of the entire population of this type of beam would be successful. Justify your response. **(2 marks)**.

7. Automobiles arrive at a vehicle equipment inspection station according to a Poisson Process with rate $\lambda = 10$ vehicles per hour. Each vehicle that arrives has a 50% chance of having no equipment violations. Assume violations on one vehicle are independent of violations on another.

a. What is the probability of exactly 10 vehicles arriving in one hour? **(2 marks)**
b. If 10 vehicles arrive what is the probability that all 10 have equipment violations? **(2 marks)**
c. What is the probability of exactly 10 vehicles arriving in one hour and all 10 having equipment violations. **(2 marks)**
d. For any fixed $y \geq 10$ what is the probability that y vehicles arrive in one hour and exactly 10 of them have equipment violations? **(2 marks)**
e. What is the probability that exactly 10 cars with equipment violations are found in one hour. **(2 marks)**

Final – ENGR 371 – August 1999

Pens, pencils, erasers, and straight edges allowed. No books. No crib sheets. Calculators allowed.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Show all your work and justify all your answers. Marks are given for how an answer is arrived at not just the answer itself.

Do ALL 7 questions given.

Total of 65 marks.

1. It may seem counterintuitive, but the distribution of the first significant digit in many collections of numbers is NOT uniform on the integers 1 through 9, but follows Benford's distribution which is as follows:

$$f(x) = \begin{cases} \log_{10}(1+\frac{1}{x}) & x = 1, 2, ...9 \\ 0 & \text{otherwise} \end{cases}$$

One use of this distribution is in auditing financial records. The idea is that if the books have been artificially altered, the distribution of the first significant digit will differ markedly from what is predicted by Benfords's distribution. Suppose that Revenue Canada is auditing the records of a large company.

   (a) If the numbers in the company's records follow Benford's distribution, what is the probability that a randomly chosen number from the company's records has a first significant digit of 1? (2 marks)

   (b) If 1000 numbers are selected randomly from the company's records and if Y is the number of those 1000 that start with 1, what is the distribution of Y? Be sure to state all parameters of the distribution. (2 marks)

   (c) If of the 1000 numbers sampled from the financial data sheets 120 started with 1 form a 95% confidence interval of the portion p of all numbers in the financial records that start with 1. (4 marks)

   (d) If of the 1000 numbers sampled from the financial data sheets 120 started with 1 would you conclude that there is evidence that the financial records have been artificially altered? Explain. (2 marks)

2. A radio signal consisting of a sinusoidal tone is sent through the air.

   (a) If you sample this signal determine the distribution of the sample value. Specifically determine the distribution of Y where $Y = \cos\theta$ where θ is uniformly distributed on $[0, 2\pi]$. (8 marks)
   The following calculus formula may be useful:

$$\frac{d}{dx}\cos^{-1} x = \frac{-1}{\sqrt{1-x^2}}$$

   (b) Sketch the pdf of Y. (2 marks)

3. Given two random variables X and Y, plot them on the $x-y$ plane. The distance from that point to the origin is $R = \sqrt{X^2 + Y^2}$. If X and Y are independent gaussian random variables both with mean 0 and standard deviation σ then R has a Rayleigh distribution which is

$$f(r) = \begin{cases} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} & r \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Determine the probability that the point lands more than 0.75σ from the origin. (3 marks)

   (b) Determine the mean of R. (Hint: Compare the equation you develop for the mean with the second moment of some gaussian distribution.) (4 marks)

   (c) Determine the probability that $2X + Y$ is more than 3σ. (3 marks)

4. Twenty rebuilt pumps were sampled from the hydraulics shop of an aircraft rework facility. The number of defects for each pump is given below.

Pump Number	1	2	3	4	5	6	7	8	9	10
No. of defects	6	3	4	0	2	7	3	1	0	0

Pump Number	11	12	13	14	15	16	17	18	19	20
No. of defects	4	3	2	2	6	5	0	7	2	1

   So pump number 12 had three defects.

   (a) Determine the sample mean of the number of defects. (2 marks)

   (b) Determine the sample median of the number of defects. (2 marks)

   (c) Determine the sample mode of the number of defects. (2 marks)

   (d) Determine the sample variance of the number of defects. (2 marks)

   (e) Determine the sample standard deviation of the number of defects. (1 marks)

   (f) Determine the sample range of the number of defects. (1 marks)

5. The arrival of phone calls at certain business is a poisson process with an average of 10 calls per hour.

   (a) What is the probability that 15 calls arrive in a certain hour? (2 marks)

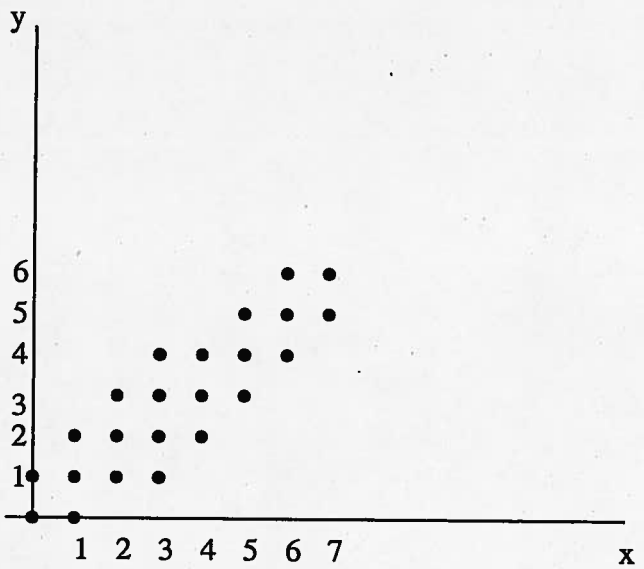
   (b) What is the standard deviation of the number of calls in one hour? (2 marks)

   (c) If the calls are routed by a receptionist in 30 seconds, what is the probability that a new call will arrive before the receptionist has time to route the current call? (2 marks)

(d) What is the mean interarrival time of calls? (2 marks)

(e) What is the standard deviation of the interrival time? (2 marks)

6. Consider two discrete random variables X and Y whose support is shown in the figure. Each of the points shown are equally likely.



(a) Determine the marginal distribution of X. (3 marks)

(b) Determine the conditional distribution of Y when X is 2. (3 marks)

(c) Are X and Y independent. Justify your response. (2 marks)

(d) Determine the probability that a point plotted on the plane according to this distribution is more than 3 units from the origin in distance. (2 marks)

7. In a paper factory paper is cut with knife blades. If the blade is "sharp" only 1% of the paper exhibits rough edges. If the blade if "average" 3% of the paper exhibits rough edges. Finally if the blade is "dull" 5% of the paper exhibits rough edges. If 25% of blades in the factory are sharp, 60% are average, and 15% are dull what is the portion of all paper that has rough edges? (5 marks)

Final – ENGR 371 – April 1999

Pens, pencils, erasers, and straight edges allowed. No books. No crib sheets. Calculators allowed.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Show all your work and justify all your answers. Marks are given for how an answer is arrived at not just the answer itself.

Do ANY 6 of the 7 questions given.

Marks: Six questions, 10 marks each. Total 60 marks.

Indicate clearly which six questions you want marked.

1. Random variable X is uniformly distributed. Specifically:

$$f(x) = \begin{cases} 0.2, & x = 0, 1, 2, 3, 4 \\ 0, & \text{otherwise} \end{cases}$$

   (a) Find the moment generating function of X.

   (b) Using the moment generating function find the first, second and third moments of X.

2. Given the joint pdf:

$$f(x,y) = \begin{cases} 2e^{-(x+y)}, & 0 \le y < \infty, \quad 0 \le x < \infty, \quad y \le x \\ 0, & \text{otherwise} \end{cases}$$

   and

$$Z = X + Y$$

   Find the pdf of Z. Note X and Y are **NOT** independent.

3. Two random variables X and Y have means $E[X] = 2$ and $E[Y] = 0$, variances $\sigma_X^2 = 1$ and $\sigma_Y^2 = 4$ and correlation coefficient $\rho_{XY} = 0.6$.

   A new random variable is defined:

$$Z = X + Y + 4$$

   Find the mean and variance of Z.

4. A controlled satellite is known to have an error (distance from target) that is normally distributed with mean zero and standard deviation 4 feet. The manufacturer of the satellite defines "success" as a firing in which the satellite comes within 10 feet of the target. Compute the probability that the satellite fails.

5. The mean breaking strength of a fabric A was found to be 25.2 pounds per square inch (psi), based on a sample of 35 specimens. The standard deviation of the sample was 5.2 psi.

   30 specimens of fabric B were also tested. The mean breaking strength was found to be 28.5 psi and the sample standard deviation was 5.9 psi.

   Assume that the populations for both fabrics are normal.

   (a) Find a 99% confidence interval for the population mean μ_A and population variance σ_A^2 for fabric A.

   (b) Find a 95% confidence interval for the difference $\mu_A - \mu_B$.

6. Two random variables X and Y have a joint distribution $f(x, y)$ with region of support given in the following figure:

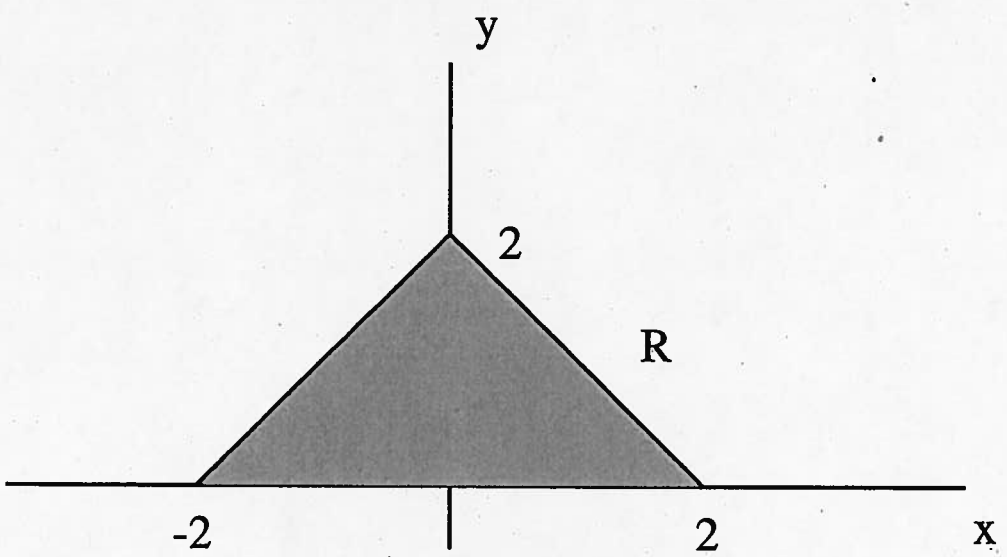


Figure 1: Region of Support

Call the region of support $\mathcal{R}$. The distribution is

$$f(x,y) = \begin{cases} K, & (x,y) \in \mathcal{R} \\ 0, & \text{otherwise} \end{cases}$$

(a) Determine K.

(b) Determine $P(X > Y)$.

(c) Determine the marginal distributions of X and Y.

(d) Determine the conditional distribution $f(x|y)$ for any value of Y.

7. The number of machine failures per day in a certain plant has a Poisson distribution with parameter $\lambda t = 3$. Present maintenance facilities can repair 3 machines per day. Failures in excess of three are repaired by a contractor.

   (a) On a given day what is the probability of having machine(s) repaired by a contractor?

   (b) If the maintenance facilities could repair four machines a day, what would be the probability of having machine(s) repaired by a contractor?

   (c) What is the expected number of machines that fail each day?

   (d) What is the expected number of machines that are repaired in the plant each day? (Hint: Be careful).

   (e) What is the expected number of machines that are repaired by the contractor each day?

Final – ENGR 371 – December 1997

Pens, pencils, erasers, and straight edges allowed. No books. No crib sheets. Calculators allowed.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Show all your work and justify all your answers. Marks are given for how an answer is arrived at not just the answer itself.

Do only **SEVEN** of the following eight questions. If you attempt all eight clearly indicate in your exam book which seven you want marked, otherwise the first seven that you attempt will be marked.

1. Two new weight loss drugs are being tested. A total of sixty subjects participate in the trials of these drugs. The first thirty take drug A and the second thirty take drug B. Performance is measured in the number of kilograms (rounded to the nearest kilogram) that the subject loses over a one month period. The results are summarized in the tables below.

Kg lost	1	2	3	4	5	6	7	8	9	10
Number	0	0	0	2	3	4	10	6	2	3

Table 1: Weight lost by population A

Kg lost	1	2	3	4	5	6	7	8	9	10
Number	0	0	0	1	0	4	10	9	3	3

Table 2: Weight lost by population B

So here 3 respondents lost 5 kg using drug A.

   (a) Compute the sample mean and the sample variance of both populations.

   (b) Compute the sample median and the sample mode, only of population A.

   (c) We wish to evaluate how much better one drug is then the other. From the first part of this question which drug do you conclude is better?

   (d) Form a 95% confidence interval on the difference of the two means, in order to judge the significance of your conclusion from the previous question.

2. Given the joint probability density function of the random variables X and Y:

$$f(x,y) = \begin{cases} kxy^3, & y \le x \le 1, 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

   (a) Calculate k such that $f(x,y)$ is a probability density function.

   (b) Find the probability that $X \le 0.5$.

3. X and Y are two independent random variables with the following distributions.

$$f(x) = \begin{cases} 1+x, & -1 < x < 0, \\ 1-x, & 0 \le x < 1, \\ 0, & \text{otherwise} \end{cases}$$

$$g(y) = \begin{cases} 0.5, & 2 < y < 4, \\ 0, & \text{otherwise} \end{cases}$$

   (a) Find the expected value of X. Call it μ_X.

   (b) Find the variance of X. Call it σ_X^2.

   (c) Define $Z = 6(2X - Y)^2 + 3X^2$. Find the expectation of Z.

   (d) Find a lower bound on $P(\mu_X - 3\sigma_X < X < \mu_X + 3\sigma_X)$.

4. Let θ be a random variable with the following distribution:

$$f(\theta) = \begin{cases} 1/2\pi, & 0 \le \theta < 2\pi, \\ 0, & \text{otherwise} \end{cases}$$

   A new random variable Y is formed using:

$$Y = A \sin\theta$$

   Here A is an unknown fixed constant. Using one value of the random variable Y we form the following estimator for A:

$$\hat{A} = k|y|$$

   Determine the constant k such that $\hat{A}$ is an unbiased estimator of A.

5. A company pays its employees an average wage of \$ 9.25/hour, with a standard deviation of 60 cents. If the wages are approximately normally distributed and paid to the nearest cent.

   (a) What percentage of the workers receive wages between \$ 8.75 and \$ 9.69 an hour inclusive?

   (b) Let S be the wage such that only 5% of the workers make more than this wage. Find S.

6. Two random variables X and Y have joint probability density function:

$$f(x,y) = \begin{cases} 8xy, & y \le x \le 1, 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

   Find the probability density function of $Z = X + Y$.

7. A word consists of 3 letters. Each letter in a word may be a vowel with probability 0.2. Whether a letter is a vowel or not is independent of the other letters. Find:

   (a) the probability that a word contains at least one vowel.

   (b) the probability that at least 12 out of 20 words will have at least one vowel.

   (c) the probability that 20 words will have a total of at least 12 vowels in them.

8. Random Processes:

   (a) Consider a random process

   $$X(t) = A\cos(\omega t + \theta)$$

   where A and θ are statistically independent random variables, and ω is a fixed frequency value. A is uniformly distributed on the range $(0, 2)$. θ is uniformly distributed on the range $(0, \pi)$.

      i. Find the mean value, the variance and the autocorrelation function of $X(t)$.

      ii. Is the process $X(t)$ Wide Sense Stationary (WSS)? Justify your answer.

   (b) Determine the mean value, and the variance of each of the random processes have the following autocorrelation functions:

      i. $R_{XX}(\tau) = 10e^{-\tau^2}$

      ii. $R_{XX}(\tau) = 10\frac{\tau^2+8}{\tau^2+4}$

Marks: 7 questions worth 10 marks each. Total 70 marks.

Final – ENGR 371 – April 1996

Pens, pencils, erasers, and straight edges allowed. No books. No crib sheets. STANDARD calculators only.

If you have a difficulty you may try making REASONABLE assumptions. State the assumption and how that assumption limits your answer. Show all your work and justify all your answers. Marks are given for how an answer is arrived at not just the answer itself.

1. You are studying traffic patterns on a single stretch of city street. On this street there is a line of vehicles. There are three vehicle types: cars, trucks and motorcycles. In a particular situation you consider 5 cars, 2 trucks and 1 motorcycle. How many distinct ways can these 8 vehicles be arranged in a line, if it is impossible to distinguish between vehicles of the same type? (4 marks)

2. In a communication system a message is sent over a noisy channel. The "message" here will be a number X between 0 and 1. Assume that all values between 0 and 1 are equally likely. i.e. assume that X has the following pdf:

$$g(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise} \end{cases}$$

The noise is a random variable N which is uniformly distributed over $[-0.5, 0.5]$. i.e. the noise N has pdf:

$$h(n) = \begin{cases} 1, & -0.5 < n < 0.5, \\ 0, & \text{otherwise} \end{cases}$$

Assume that the message and the noise are independent. The noise is additive. That is, the message after corruption is

$$Y \triangleq X + N$$

   (a) Write down the joint distribution of X and N. (2 marks)

   (b) Determine the mean of X. (2 marks)

   (c) Determine the variance of X. (2 marks)

   (d) Determine the distribution of Y. (7 marks)

   (e) Determine the mean of Y. (2 marks)

   (f) Determine the variance of Y. (2 marks)

   (g) What is $P(|X - Y| > 0.25)$? (4 marks)

3. Suppose that for a very large shipment of integrated chips the probability of failure for any one chip is 10%. Taking a sample of 20 chips what is the probability that 3 or fewer of the chips are defective. Assume that whether one chip fails or not is independent of the failures of other chips in the sample. (5 marks)

4. Consider that customers arrive in a store according to a Poisson Process, with rate parameter $\lambda = 5$ customers/hour. The store is open from 9 am to 8 pm.

   (a) What is the probability that no customers arrive in the first hour (from 9 am to 10 am)? (3 marks)

   (b) What is the expected number of customers that arrive when the boss is at lunch (between 12 and 12:30)? (2 marks)

   (c) Let U be the number of customers that arrive between 1 pm and 2 pm. Let V be the number of customers that arrive between 2pm and 3 pm. Write down the conditional distribution $f(V|U = 5)$. (4 marks)

   Assume that on a certain day ther are at least 3 customers. Let X be the waiting time between the first and second customer. Let Y be the waiting time between the second and third customers.

   (d) Write down the distribution of X. (2 marks)

   (e) Write down the distribution of Y. (2 marks)

   (f) Write down the joint distribution of X and Y. (2 marks)

   (g) Determine $P(X > 3Y)$. (4 marks)

5. Monte Carlo Integration: Monte Carlo integration is used to evaluate difficult integrals. One is given a function $f(x)$ which is difficult to express analytically. However we can tell whether a point (x, y) is above or below $f(x)$. For this question assume that $f(x)$ is non-negative everywhere. To evaluate:

$$I = \int_b^a f(x)dx$$

   We choose a d such that $d > f(x), \forall x \in [b, a]$. Then we generate random points uniformly distributed over the rectangle:

$$b < x < a$$

$$0 < y < d$$

   (To do this generate a random variable X uniformly distributed on $[b, a]$ and another random variable Y uniformly distributed on $[0, d]$. They should be independent of each other. Then (X, Y) is uniformly distributed on the rectangle. See Figure 1.)

   The area of the rectangle, call it A, is clearly $A = d * (a - b)$.

   The following will help us decide how we can estimate I.

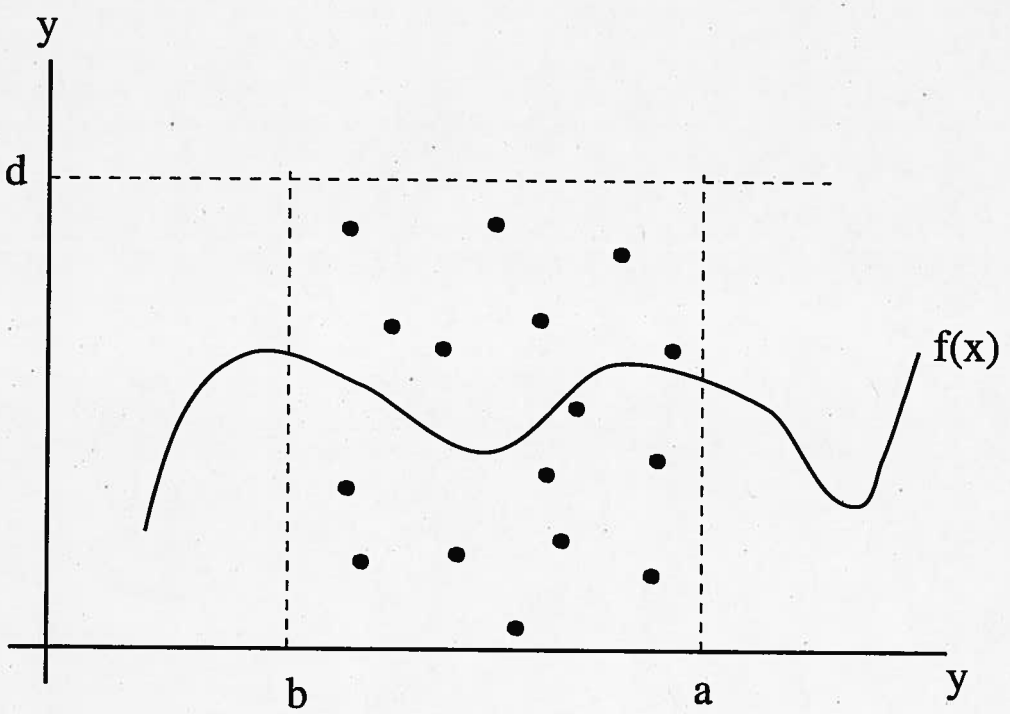


Figure 1: Monte Carlo Integration

(a) Consider one point uniformly distributed on the rectangle. What is the probability that it is below $f(x)$. State your answer in terms of A and I. (2 marks)

(b) Consider 1000 points uniformly distributed on the rectangle. Each is independent of the others. For the i'th point (X_i, Y_i) define a R.V. U_i such that:

$$U_i = \begin{cases} 1, & (X_i, Y_i) \text{ below } f(X_i) \\ 0, & \text{otherwise} \end{cases}$$

Give a pmf for U_i. (3 marks)

(c) Define:

$$T = \sum_{i=1}^{1000} U_i$$

What kind of distribution does T have? Be sure to specify the parameters for that distribution and the values that those parameters take on here. (4 marks)

(d) You should NOT have gotten a gaussian distribution for the last question. However the distribution for T can be approximated as a gaussian, perhaps under certain conditions. JUSTIFY this statement, giving any conditions that are necessary. Then specify which gaussian by giving the parameters for the gaussian and the values of those parameters in this case. (3 marks)

(e) Consider the following estimator for I

$$\hat{I} = A\frac{T}{1000}$$

Is this an unbiased estimate of I? Justify your response. (2 marks)

Bonus Question

(f) Form a 95% confidence interval for I. (4 bonus marks)

6. We wish to draw inferences about a population. To do this, we obtain a random sample. The outcomes of this random sample are shown in the following table:

X_i	100	200	300	400	500	600	700
f_i	16	24	12	10	8	5	2

Here the samples X_i took only values that were multiples of 100. f_i is the number of times each of the possible values of x_i were observed. For example the outcome 200 was observed 24 times in this random sample.

(a) Compute the sample mean and the sample variance. (4 marks)

(b) Compute the sample median and the sample mode. (3 marks)

(c) Form a 95% confidence interval of the mean of the population. (4 marks)

7. Consider a discrete parameter strictly stationary, ergodic, zero mean gaussian random process $X(n)$ with the following correlation function:

$$R_{xx}(\tau) = \begin{cases} 1, & \tau = 0, \\ 0, & \text{otherwise} \end{cases}$$

(a) Determine the value E[X(1) X(5)]. (2 marks)

(b) Determine the value E[X(2) X(6)]. (2 marks)

(c) What is the variance of $X(3)$? (2 marks)

Total 80 marks plus 4 Bonus marks

Some Useful Equations

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$$

$$\sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1 - r}$$