STATISTICS

Welcome to Math 3200! My name is Professor Edward Spitznagel. This is the successor course to Math 320. It is a calculus-based introductory course in statistics and the underlying probability theory supporting it. Since this course is now differentiated (and integrated—☺) from the effectively non-calculus-based Math 2200, a paragraph or two of explanation is warranted.

When I began teaching Math 320 in 1970, it had an enrollment of 21 students. At that time, it was a calculus-based course. Over the years, it grew until a few years ago it had over 400 students. Gradually, the calculus prerequisite became a nominal one-semester dose (Math 131), which meant that the quality of the course really suffered. Perhaps that would not have been a problem, except for the fact that many of our upper level courses depended on students being prepared for them by Math 320. Without that preparation, Those courses had to spend their first third in reviewing what should have been covered in Math 320, and thus themselves became watered down.

By returning Math 320 to its roots, we hope to upgrade the quality of all our statistics offerings for both mathematics majors and minors. Of course, any student, major, minor, or not, who has the calculus background is welcome in the revitalized Math 320. Although what we are doing is in fact restoring Math 320 to what it once was, it was decided that it might be more politically correct to give it a new number—thus its new designation as Math 3200.

Times and Places

Our course meets Monday, Wednesday, and Friday 9-10 in Brown 118. **Before you come to class, please preview the section(s) of the book to be covered that day.** Naturally I don't expect you to learn all the material from that reading. What I do expect is that you will be able to ask much better questions, having done that preview. You can tell which sections to have a look at from the homework problems assigned on Pages 4-5 of this syllabus

My *official* office hours are from 8:00 to 8:45 on Monday, Wednesday, and Friday in Brown 118. After my teaching ends at noon, I typically take the Metro down to the Med School. You might find me in Room 118 of Cupples I when I get back late in the afternoon. You are *welcome* to knock anytime you see the light on. However, I do recommend calling in advance to see if I'm in. My telephone number is 935-6745.

Textbook

The text is Tamhane and Dunlop *Statistics and Data Analysis: From Elementary to Intermediate*. This is one of a very few books from which a junior level course can be taught. Most other books are either too hard (too much mathematics) or too soft (too little mathematics). Like Baby Bear's bed, Tamhane and Dunlop is just right. I have to confess that I did use the book once before, for all of the previous Math 320, and there was a lot of kvetching about it. It seems to have been a matter of ***μαργαριτας εμπροσθεν χοιρον,*** which I don't think will apply here.

Hand Held Technology

The **Texas Instruments calculators TI-83, TI-84, and TI-89 (and the new TI-Nspire series)** contain essentially every probability function and statistical program we will be using during the course. I have declared the above to be the official calculators for the course. I have a computer emulation of the TI-83, with which I will frequently work problems in class, projecting an image of the calculator on the screen. These calculators contain probability functions that are more accurate and easier to use than the distribution tables in the back of the book. Therefore I will not provide those tables for the examinations; you will be expected to use the calculator instead. ***Verbum sapienti!***

Manual Homework

There are six recommended homework problems per day of class. In most instances, two are odd-numbered, with answers in the back of the book. The other four are even-numbered book problems or are taken from the "A" questions listed at the end of this syllabus. I will usually have time to work two even-numbered problems in class, leaving you with a net four problems per day to do on your own. These problems will not be graded. Your primary motivation for keeping up with the homework is that most of the examination problems will be homework problems with simple changes in the data.

For those of you who wish it, a grader will provide you with feedback via email on any problems you choose to do. Those who participated regularly in this service last year all achieved course grades of A– or higher. By 9AM of the Tuesdays and Thursdays following the Monday and Wednesday classes, you may drop off your solutions of whatever problems you wish in the Math Dept office, Room 100 of Cupples I. Following the Friday class, you may drop your solutions in the Room 100 door's mail slot by noon Saturday.

Please write only on the front side of each page, use a paperclip (not a staple) to hold them together, and pull off any jaggies if you tore them out of a notebook. Print the course number (Math 3200) and your email address ***clearly*** at the top of each page. We will score your solutions and email you scanned copies.

For those of you studying as a team, just submit one copy. Whoever submits it will receive the email and can forward it to everyone else. We're sorry that, due to the limitations of our scanner, we can only email a scored assignment back to a single address.

There are three simple conditions to this offer. First, we will only score original, handwritten work, not photocopies. Second, we will only score good-faith attempts to solve the problems; we will not write in solutions, or even provide answers, on blank sheets of paper. Third, we will not score illegible solutions; we will simply return these marked as illegible.

We will keep no records of how well you did on these problems. This is strictly a feedback service. There is no need to give us your name; just provide your email address.

Computer Technology

There is a wide variety of computer software for doing statistics, ranging from the relatively primitive capabilities in Microsoft Excel® to the extremely powerful SAS® package. We will use four statistics packages: SAS, STATA®, R®, and SPSS®. I introduce you to these four main packages because all are in common use, and you can never tell which one(s) might get you that envied job or internship. Last year a student wrote me the following thank-you note: "I'm starting an internship with the data analytic department at _____ next semester, and they were extremely impressed that I already had an introduction to SAS and SPSS. I don't think I would even have had a chance at that internship if I didn't take the class."

I will demonstrate all four in class, and will assign homework problems for you to do and hand in for grading. The primary package will be SAS. We will cover the others in compare-and-contrast mode, so that you will be able to claim at least passing familiarity with all four when the time comes to interview for jobs and internships. The ArtSci laboratory in the basement of Seigle Hall has SAS, STATA, and R installed on its computers. When the time comes, I will introduce you to SPSS via a freeware clone called PSPP.

Computer Homework

There are three required computer homework exercises per week of class. When it is convenient, these problems are chosen from the recommended manual homework problems. These exercises are due in class each Monday, with the exception of Labor day and the Mondays immediately following Fall Break and Thanksgiving. That works out to a total of ten assignments. The computer homework will count as 20% of your course grade.

Unfortunately, the number of computers publicly available to Arts and Science undergraduates has dwindled from more than 60 down to the 14 currently in Seigle Hall Room L012: 10 PC's and 4 iMac's. Prior to this year, I used to drop by Sunday afternoons to the large and spacious lab beneath Holmes Lounge. Typically more than half the class would show up at that time to do their homework. My eagle eye would help them spot missing semicolons, unbalanced quotes, and other little details that can drive a person to drink. Regrettably, with the small size of the Seigle Lab, that is no longer feasible.

Thankfully, though, Seigle L012 is open generous hours:

Monday:	8:30am-8:30pm
Tuesday:	8:30am-8:30pm
Wednesday:	8:30am-8:30pm
Thursday:	8:30am-8:30pm
Friday:	8:30am-5:30pm
Saturday:	11:00am-5:00pm
Sunday:	11:00am-5:00pm

As long as you don't all wait until late Sunday afternoon, you should be ok. As far as dealing with those mischievous quotation marks and semicolons, feel free to bring snippets of code (printed on paper or copied to flash drive) to my office hours. I will have the appropriate software running on the computer for you to show me what is bugging you. Also, if you get the feeling that 14 public computers are insufficient for 3955 A&S undergraduates, you might mention that to your academic advisors.

Examinations

As mentioned earlier, examinations are closely linked to the homework problems. If you faithfully work the problems, you should have no trouble scoring well on the examinations. Each examination will contain twenty multiple choice problems, of which approximately fifteen will be homework problems with altered numbers. You may bring one 4×6 inch notecard to each in-semester examination, and up to four notecards to the final examination. For the final, students usually recycle the first three cards and put the new stuff on a fourth one. You may use both sides of each notecard.

Over the four examinations, you can achieve a maximum of 80 points. With the computer homework added in, your maximum number of points will be 100. At the end of the semester, the A range will be 90 and above, the B range will be 80 to 90, the C range will be 70 to 80, and the D range will be 60 to 70, with plus and minus grades at the tops and bottoms of each of these ranges.

Students ask if I ever grade on a "curve." Curve grading was popular about fifty years ago. It assigned six letter grades A, B, C, D, E, and F based on a Gaussian, also called a "normal" curve. The grade of A corresponded to being 2 standard deviations above the mean and was awarded to the upper 2.5% of all students. The grade of B corresponded to being one to two standard deviations above the mean and was awarded to 13.6% of all students. The most common grades were C and D, at 34.1% each. I doubt any of you would like the grades to be assigned based on that system.

Instead, I will follow the modern convention, in which the A range will be 90 to 100, the B range will be 80 to 90, the C range will be 70 to 80, and the D range will be 60 to 70, with plus and minus grades at the tops and bottoms of each of these ranges. If you are registered pass/fail, you must achieve at least 70 points to pass, which is the lowest score for a C−.)

In addition to calculating the straight sum of points, I will also average the examination scores using a weighting process, in which each in-semester examination counts 16% and the final counts 32%, giving you whichever score is higher. (The computer homework will still be counted at 20%.)

This alternative weighting system rewards students who have tended to improve over the semester.

Examination Schedule

The three in-semester examinations will be given from 7PM to 9PM the following **Wednesday evenings**: September 19th, October 17th, and November 14th.

The final examination will be given on **Thursday, December 13, 3:30PM-5:30PM**.

As always, examination room assignments are posted on the Math Dept website:

http://www.math.wustl.edu/seatlookup/

the day of the examination.

Recommended Homework

Following are the recommended homework problems. At the risk of preaching to the choir, let me say that mastering these and reading the book should give you the traditional two-hours-out-of-class-for-every-one-in-class needed for success in the typical undergraduate course. The last time I taught a course to engineering students, they complained to their dean that I was working them too hard, giving them homework that took two hours (gasp, shudder) per class period. He asked them how much homework time they spent in their other courses, and they said, oh, about half an hour. Would you really want to fly in an airplane designed by engineers like that?

Two schools, CalTech and MIT, award credits equal to the weekly sum of lecture hours and expected amount of hours outside of class. As a reality check, I visited their websites and found the credits for their equivalent statistics courses to be:

CalTech:	Ma112a lists 9 units of credit.
MIT:	18.443 lists 12 units of credit.

Thus, these two schools expect their students to spend between two and three hours outside of class for every hour inside class.

Aug 29	Chapter 2	6,17,26,27,A1,A3
Aug 31	Chapter 2	29,34,35,40,A5,A7
Sept 3	**Labor Day Holiday**	
Sept 5	Chapter 2	50,54,55,57,A10,A43
Sept 7	Chapter 2	59,61,62,64,A32,A38
Sept 10	Chapter 2	72,81,83,A42,A44,A51
Sept 12	Chapter 3	2,3,8,10,12,15
Sept 14	Chapter 3	19,20,22,23,24,26
Sept 17	Chapter 4	4,12,13,33,38,40
Sept 19	Chapter 5	8,9,17,24,30,36
Sept 19	**First Examination, 7-9PM**	
Sept 21	Chapter 6	2,4,5,7,12,14
Sept 24	Chapter 6	18,19,20,22,27,30
Sept 26	Chapter 7	2,6,7,13,16,18
Sept 28	Chapter 8	3,4,5,6,7,8
Oct 1	Chapter 8	9,12,14,17,18,20
Oct 3	Chapter 9	2,6,12,13,15,16
Oct 5	Chapter 9	18,19,20,28,29,36
Oct 8	Chapter 10	2,4,5,6,7,8
Oct 10	Chapter 10	11,12,13,20,24,26
Oct 12	Chapter 10	28,29,30,32,34,36
Oct 15	Chapter 11	2,3,4,5,6,10
Oct 17	Chapter 11	11,12,14,15,17,18
Oct 17	**Second Examination, 7-9PM**	
Oct 19	**Fall Break**	
Oct 22	Chapter 11	21,22,23,24,26,28
Oct 24	Chapter 11	30,31,34,36,38,39

Oct 26	Chapter 11	40,42,44,45,46,49
Oct 29	Chapter 12	1,2,3,4,5,6
Oct 31	Chapter 12.	8,10,11,12,15,16
Nov 2	Chapter 12	18,20,21,22,24,27
Nov 5	Chapter 13	1,2,4,5,10,14
Nov 7	Chapter 13	16,17,18,19,20,22
Nov 9	Chapter 13	24,26,27,28,29,30
Nov 12	Chapter 14	2,4,5,8,9,10
Nov 14	Chapter 14	12,14,16,19,20,21
Nov 14	**Third Examination, 7-9 PM**	
Nov 16	Chapter 14	26,28,29,34,36,37
Nov 19	Chapter 15	1,2,3,6,11,12
Nov 21	**Thanksgiving Holiday**	
Nov 23	**Thanksgiving Holiday**	
Nov 26	Chapters 15/8	13,14,15,17,8.1*,8.2*
Nov 28	Chapters 15/12	12.2*
Nov 30	Chapters 15/12	12.1*,12.3*

*The starred problems come from Chapters 8 and 12 of an e-book downloadable (for free) from Olin Library:
`http://dx.doi.org/10.1201/9781420057553`

Dec 3	Chapter 15	18,19,22,23
Dec 5	Chapter 15	24,27,29,31
Dec 7	Chapter 15	33,35
Dec 10-12	**Reading Period**	
Dec 13	**Final Examination, 3:30-5:30 PM**	

Required Homework

Here are the required computer homework problems. Three problems are due per week, **always on Monday, at the beginning of class**. Three Mondays are skipped, making the total number of assignments equal to eleven. All assignments are to be done with SAS. In addition, most assignments are also to be done with STATA, R, or SPSS, on a semi-rotating basis. I will let you know on a week-by-week basis what other package is to be used that week, in addition to SAS.

Sept 10	2.26 (use proc freq with n=10000 and use the table percentages to answer the questions), 2.29b (simulate 100000 times), A43 (simulate 100000 times)
Sept 17	2.81, A44, A51 (simulate each one 100000 times),
Sept 24	4.12, 5.9, 5.24
Oct 1	6.2 (simulate 100000 times), 6.7 (use proc means), 7.16 (use proc means)
Oct 8	8.12bc, 8.17, 9.13 (use proc freq to perform the Fisher exact test)
Oct 15	10.4, 10.11, 10.36
Oct 29	11.6, 11.22, 11.45
Nov 5	12.5, 12.10, 12.20
Nov 12	13.2, 13.14, 13.26
Nov 19	14.20, 14.26, 14.36 (draw 1000 bootstrap samples, rather than just 25)
Dec 3	15.17, 12.1*, 12.2* (the latter two problems are from Chapter 8 of the e-book)

"A" Problems

A1. A survey of a group's viewing habits over the last year revealed the following information:

(i) 28% watched gymnastics
(ii) 29% watched baseball
(iii) 19% watched soccer
(iv) 14% watched gymnastics and baseball
(v) 12% watched baseball and soccer
(vi) 10% watched gymnastics and soccer
(vii) 8% watched all three sports.

Calculate the percentage of the group that watched none of the three sports during the last year.

A2. The probability that a visit to a primary care physician's (PCP) office results in neither lab work nor referral to a specialist is 35%. Of those coming to a PCP's office, 30% are referred to specialists and 40% require lab work. Determine the probability that a visit to a PCP's office results in both lab work and referral to a specialist.

A3. A public health researcher examines the medical records of a group of 937 men who died in 1999 and discovers that 210 of the men died from causes related to heart disease. Moreover, 312 of the 937 men had at least one parent who suffered from heart disease, and, of these 312 men, 102 died from causes related to heart disease. Determine the probability that a man randomly selected from this group died of causes related to heart disease, given that neither of his parents suffered from heart disease.

A4. An insurance company examines its pool of auto insurance customers and gathers the following information:

(i) All customers insure at least one car.
(ii) 70% of the customers insure more than one car.
(iii) 20% of the customers insure a sports car.
(iv) Of those customers who insure more than one car, 15% insure a sports car.

Calculate the probability that a randomly selected customer insures exactly one car and that car is not a sports car.

A5. An insurance company determines that N, the number of claims received in a week, is a random variable with $P(N = n) = 1/(2^{n+1})$, where $n \geq 0$. The company also determines that the number of claims received in a given week is independent of the number of claims received in any other week. Determine the probability that exactly seven claims will be received during a given two-week period.

A6. The loss due to a fire in a commercial building is modeled by a random variable X with density function $f(x) = 0.005(20-x)$ for $0<x<20$, and 0 elsewhere. Given that a fire loss exceeds 8, what is the probability that it exceeds 16?

A7. An insurance policy pays an individual \$1000 per day for up to 3 days of hospitalization and \$250 per day for each day of hospitalization thereafter. The number of days of hospitalization, X, is a discrete random variable with p.m.f. = $(6-x)/15$ for $x=1,2,3,4,5$. Calculate the expected payment for hospitalization under this policy.

A8. An actuary determines that the claim size for a certain class of accidents is a random variable X with moment generating function

$$M_X(t) = (1/(1-2500t))^4$$

Determine the standard deviation of the claim size for this class of accidents.

A9. An insurance policy pays a total medical benefit consisting of two parts for each claim. Let X represent the part of the benefit that is paid to the surgeon, and let Y represent the part that is paid to the hospital. The variance of X is 5000, the variance of Y is 10000, and the variance of the total benefit, X+Y , is 17000. Due to increasing medical costs, the company that issues the policy decides to increase X by a flat amount of 100 per claim and to increase Y by 10% per claim. Calculate the variance of the total benefit after these revisions have been made.

A10. A company insures homes in three cities, J, K, and L. Since sufficient distance separates the cities, it is reasonable to assume that the losses occurring in these cities are independent. The moment generating functions for the loss distributions of the cities are:

$M_J(t) = (1-2t)^{-3}$
$M_K(t) = (1-2t)^{-2.5}$
$M_L(t) = (1-2t)^{-4.5}$

Let X represent the combined losses from the three cities.
Calculate $E(X^3)$.

A11. An actuary has discovered that policyholders are three times as likely to file two claims as to file four claims. If the number of claims filed has a Poisson distribution, what is the variance of the number of claims filed?

A12. A company establishes a fund of \$12,000 from which it wants to pay a bonus to any of its 20 employees who achieve a high performance level during the coming year. Each employee has a 2% chance of achieving a high performance level during the coming year, independent of any other employee. Determine the maximum value of the bonus for which the probability is less than 1% that the fund will be inadequate to cover all payments for high performance.

A13. The number of days that elapse between the beginning of a calendar year and the moment a high-risk driver is involved in an accident is exponentially distributed. An insurance company expects that 30% of high-risk drivers will be involved in an accident during the first 50 days of a calendar year. What portion of high-risk drivers are expected to be involved in an accident during the first 80 days of a calendar year?

A14. The lifetime of a printer costing 200 is exponentially distributed with mean 2 years. The manufacturer agrees to pay a full refund to a buyer if the printer fails during the first year following its purchase, and a one-half refund if it fails during the second year. If the manufacturer sells 100 printers, how much should it expect to pay in refunds?

A15. Two instruments are used to measure the height, h, of a tower. The error made by the less accurate instrument is normally distributed with mean 0 and standard deviation 0.0056h. The error made by the more accurate instrument is normally distributed with mean 0 and standard deviation 0.0044h. Assuming the two measurements are independent random variables, what is the probability that their average value is within 0.005h of the height of the tower?

A16. A company manufactures a brand of light bulb with a lifetime in months that is normally distributed with mean 3 and variance 1 . A consumer buys a number of these bulbs with the intention of replacing them successively as they burn out. The light bulbs have independent lifetimes. What is the smallest number of bulbs to be purchased so that the succession of light bulbs produces light for at least 40 months with probability at least 0.9772?

A17. Claims filed under auto insurance policies follow a normal distribution with mean 19,400 and standard deviation 5,000. What is the probability that the average of 25 randomly selected claims exceeds 20,000?

A18. For Company A there is a 60% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 10,000 and standard deviation 2,000. For Company B there is a 70% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 9,000 and standard deviation 2,000. Assume that the total claim amounts of the two companies are independent. What is the probability that, in the coming year, Company B’s total claim amount will exceed Company A’s total claim amount?

A19. A person is dealt a 13-card bridge hand. Contrary to the rules, he shouts out, “Oh goody, I have an ace!” What is the probability that in fact he has more than one ace?

A20. A person is dealt a 13-card bridge hand. Contrary to the rules, he shouts out, “Oh goody, I have the ace of spades!” What is the probability that in fact he has more than one ace?

A21. A doctor is studying the relationship between blood pressure and heartbeat abnormalities in her patients. She tests a random sample of her patients and notes their blood pressures (high, low, or normal) and their heartbeats (regular or irregular). She finds that:

(i) 14% have high blood pressure.
(ii) 22% have low blood pressure.
(iii) 15% have an irregular heartbeat.
(iv) Of those with an irregular heartbeat, one-third have high blood pressure.
(v) Of those with normal blood pressure, one-eighth have an irregular heartbeat.

What portion of the patients selected have a regular heartbeat and low blood pressure?

A22. An actuary is studying the prevalence of three health risk factors, denoted by A, B, and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other). The probability that a woman has all three risk factors, given that she has A and B, is 1/3. What is the probability that a woman has none of the three risk factors, given that she does not have risk factor A?

A23. An insurance company issues life insurance policies in three separate categories: standard, preferred, and ultra-preferred. Of the company's policyholders, 50% are standard, 40% are preferred, and 10% are ultra-preferred. Each standard policyholder has probability 0.010 of dying in the next year, each preferred policyholder has probability 0.005 of dying in the next year, and each ultra-preferred policyholder has probability 0.001 of dying in the next year. A policyholder dies in the next year. What is the probability that the deceased policyholder was ultra-preferred?

A24. A health study tracked a group of persons for five years. At the beginning of the study, 20% were classified as heavy smokers, 30% as light smokers, and 50% as nonsmokers. Results of the study showed that light smokers were twice as likely as nonsmokers to die during the five-year study, but only half as likely as heavy smokers. A randomly selected participant from the study died over the five-year period. Calculate the probability that the participant was a heavy smoker.

A25. The number of injury claims per month is modeled by a random variable N with $P[N = n] = 1/((n+1)(n+2))$, where $n \geq 0$. Determine the probability of at least one claim during a particular month, given that there have been at most four claims during that month.

A26. A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease. Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

A27. A hospital receives 1/5 of its flu vaccine shipments from Company X and the remainder of its shipments from other companies. Each shipment contains a very large number of vaccine vials. For Company X's shipments, 10% of the vials are ineffective. For every other company, 2% of the vials are ineffective. The hospital tests 30 randomly selected vials from a shipment and finds that one vial is ineffective. What is the probability that this shipment came from Company X?

A28. An insurance company insures a large number of homes. The insured value, X, of a randomly selected home is assumed to follow a distribution with density function $f(x) = 3x^{-4}$ for $x>1$ and $f(x) = 0$ otherwise. Given that a randomly selected home is insured for at least 1.5, what is the probability that it is insured for less than 2?

A29. A large pool of adults earning their first driver's license includes 50% low-risk drivers, 30% moderate-risk drivers, and 20% high-risk drivers. Because these drivers have no prior driving record, an insurance company considers each driver to be randomly selected from the pool. This month, the insurance company writes 4 new policies for adults earning their first driver's license. What is the probability that these 4 will contain at least two more high-risk drivers than low-risk drivers?

A30. The lifetime of a machine part has a continuous distribution on the interval (0, 40) with probability density function f, where f(x) is proportional to $1/(10+x)^2$. What is the probability that the lifetime of the machine part is less than 5?

A31. A company prices its hurricane insurance using the following assumptions:

(i) In any calendar year, there can be at most one hurricane.
(ii) In any calendar year, the probability of a hurricane is 0.05 .
(iii) The number of hurricanes in any calendar year is independent of the number of hurricanes in any other calendar year.

Using the company's assumptions, calculate the probability that there are fewer than 3 hurricanes in a 20-year period.

A32. A study is being conducted in which the health of two independent groups of ten policyholders is being monitored over a one-year period of time. Individual participants in the study drop out before the end of the study with probability 0.2 (independently of the other participants). What is the probability that at least 9 participants complete the study in one of the two groups, but not in both groups?

A33. A company takes out an insurance policy to cover accidents that occur at its manufacturing plant. The probability that one or more accidents will occur during any given month is 3/5. The number of accidents that occur in any given month is independent of the number of accidents that occur in all other months. Calculate the probability that there will be at least four months in which no accidents occur before the fourth month in which at least one accident occurs.

A34. Let X be a continuous random variable with density function $f(x) = |x|/10$ for $-2 \le x \le 4$ and $f(x)=0$ otherwise. Calculate the expected value of x.

A35. A device that continuously measures and records seismic activity is placed in a remote region. The time T to failure of this device is exponentially distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T,2)$. Determine $E[X]$.

A36. A piece of equipment is being insured against early failure. The time from purchase until failure of the equipment is exponentially distributed with mean 10 years. The insurance will pay an amount x if the equipment fails during the first year, and it will pay 0.5x if failure occurs during the second or third year. If failure occurs after the first three years, no payment will be made. At what level must x be set if the expected payment made under this insurance is to be $1000?

A37. An insurance policy on an electrical device pays a benefit of $4000 if the device fails during the first year. The amount of the benefit decreases by $1000 each successive year until it reaches 0. If the device has not failed by the beginning of any given year, the probability of failure during that year is 0.4. What is the expected benefit under this policy?

A38. A company buys a policy to insure its revenue in the event of major snowstorms that shut down business. The policy pays nothing for the first such snowstorm of the year and $10,000 for each one thereafter, until the end of the year. The number of major snowstorms per year that shut down business is assumed to have a Poisson distribution with mean 1.5. What is the expected amount paid to the company under this policy during a one-year period?

A39. An insurance policy reimburses a loss up to a benefit limit of 10. The policyholder's loss, Y, follows a distribution with density function $f(y) = 2/y^3$ for $y>1$, $f(y) = 0$ otherwise. What is the expected value of the benefit paid under the insurance policy?

A40. The warranty on a machine specifies that it will be replaced at failure or age 4, whichever occurs first. The machine's age at failure, X, has density function $f(x) = 1/5$ for $0<x<5$, $f(x) = 0$ otherwise. Let Y be the age of the machine at the time of replacement. Determine the variance of Y.

A41. A baseball team has scheduled its opening game for April 1. If it rains on April 1, the game is postponed and will be played on the next day that it does not rain. The team purchases insurance against rain. The policy will pay 1000 for each day, up to 2 days, that the opening game is postponed. The insurance company determines that the number of consecutive days of rain beginning on April 1 is a Poisson random variable with mean 0.6 . What is the standard deviation of the amount the insurance company will have to pay?

A42. The time to failure of a component in an electronic device has an exponential distribution with a median of four hours. Calculate the probability that the component will work without failing for at least five hours.

A43. A device runs until either of two components fails, at which point the device stops running. The joint density function of the lifetimes of the two components, both measured in hours, is $f(x,y) = (x+y)/27$ for $0<x<3$ and $0<y<3$. Calculate the probability that the device fails during its first hour of operation.

A44. A charity receives 2025 contributions. Contributions are assumed to be independent and identically distributed with mean 3125 and standard deviation 250. Calculate the approximate 90th percentile for the distribution of the total contributions received.

A45. Claims filed under auto insurance policies follow a normal distribution with mean \$19,400 and standard deviation \$5,000. What is the probability that the average of 25 randomly selected claims exceeds \$20,000?

A46. In an analysis of healthcare data, ages have been rounded to the nearest multiple of 5 years. The difference between the true age and the rounded age is assumed to be uniformly distributed on the interval from –2.5 years to 2.5 years. The healthcare data are based on a random sample of 48 people. What is the approximate probability that the mean of the rounded ages is within 0.25 years of the mean of the true ages?

A47. The waiting time for the first claim from a good driver and the waiting time for the first claim from a bad driver are independent and follow exponential distributions with means 6 years and 3 years, respectively. What is the probability that the first claim from a good driver will be filed within 3 years and the first claim from a bad driver will be filed within 2 years?

A48. Two insurers provide bids on an insurance policy to a large company. The bids must be between 2000 and 2200. The company decides to accept the lower bid if the two bids differ by 20 or more. Otherwise, the company will consider the two bids further. Assume that the two bids are independent and are both uniformly distributed on the interval from 2000 to 2200. Determine the probability that the company considers the two bids further.

A49. A tour operator has a bus that can accommodate 20 tourists. The operator knows that tourists may not show up, so he sells 21 tickets. The probability that an individual tourist will not show up is 0.02, independent of all other tourists. Each ticket costs \$50, and is non-refundable if a tourist fails to show up. If a tourist shows up and a seat is not available, the tour operator has to pay \$100 (ticket cost + \$50 penalty) to the tourist. What is the expected revenue of the tour operator?

A50. Let X_1, X_2, X_3 be a random sample from a discrete distribution with probability function $p(0) = 1/3$, $p(1) = 2/3$, $p(x) = 0$ otherwise. Determine the moment generating function, $M_Y(t)$, of $Y = X_1X_2X_3$.

A51. A company has two electric generators. The time until failure for each generator follows an exponential distribution with mean 10. The company will begin using the second generator immediately after the first one fails. What is the variance of the total time that the generators produce electricity?

A52. A joint density function is given by $f(x,y) = kx$, $0<x<1$, $0<y<1$, $f(x,y) = 0$ otherwise, where k is a constant. What is $COV(X,Y)$?

A53. The value of a piece of factory equipment after three years of use is $100(0.5)^X$ where X is a random variable having moment generating function $M_X(t) = 1/(1-2t)$ for $t<1/2$. Calculate the expected value of this piece of equipment after three years of use.

A54. A store has 80 modems in its inventory, 30 coming from Source A and the remainder from Source B. Of the modems from Source A, 20% are defective. Of the modems from Source B, 8% are defective. Calculate the probability that exactly two out of a random sample of five modems from the store's inventory are defective.

A55. The number of workplace injuries, N, occurring in a factory on any given day is Poisson distributed with mean λ. The parameter λ is a random variable that is determined by the level of activity in the factory, and is uniformly distributed on the interval $[0,3]$. Calculate $Var(N)$.

A56. A survey of 100 TV watchers revealed that over the last year:
- i) 34 watched CBS.
- ii) 15 watched NBC.
- iii) 10 watched ABC.
- iv) 7 watched CBS and NBC.
- v) 6 watched CBS and ABC.
- vi) 5 watched NBC and ABC.
- vii) 4 watched CBS, NBC, and ABC.
- viii) 18 watched HGTV and of these, none watched CBS, NBC, or ABC.

Calculate how many of the 100 TV watchers did not watch any of the four channels (CBS, NBC, ABC or HGTV).

A57. The amount of a claim that a car insurance company pays out follows an exponential distribution. By imposing a deductible, the insurance company reduces the expected claim payment by 10%. Calculate the percentage reduction on the variance of the claim payment.

A58. From 27 pieces of luggage, an airline luggage handler damages a random sample of four. The probability that exactly one of the damaged pieces of luggage is insured is twice the probability that none of the damaged pieces are insured. Calculate the probability that exactly two of the four damaged pieces are insured.

A59. Let X represent the number of customers arriving during the morning hours and let Y represent the number of customers arriving during the afternoon hours at a diner. You are given:
- i) X and Y are Poisson distributed.
- ii) The first moment of X is less than the first moment of Y by 8.
- iii) The second moment of X is 60% of the second moment of Y.

Calculate the variance of Y.

A60. According to legend, Wild Bill Hickok was holding a very specific five-card poker hand when he was gunned down by Jack McCall in a Deadwood, South Dakota, saloon. The hand contained a pair of aces, a pair of eights, and a fifth card that was neither ace nor eight. What is the probability of being dealt this “Dead Man’s Hand” (also called “Aces and Eights”)? What is the probability of being dealt the *exact* cards shown in the logo below (black aces, black eights, and the jack of diamonds)?