Stat 5351 Probability & Statistics I

Tentative Exam dates
Exam 1: Sept. 29, 2011
Exam 2: Nov. 8, 2011
Exam 3: 5:00-7:45 pm, Dec. 10, 2011

Syllabus

Stat 5351 Course Information

Instructor:	Dr. Larry P. Ammann
Office hours:	T-Th 4-5pm, others by appt.
Email:	ammann@utdallas.edu
Office:	FO 2.904D
Phone:	(972) 883-2162
Text:	Mathematical Statistics with Applications, 7^{rd} Edition
Authors:	Miller and Miller

Topics	**Chapters**
Introduction	class notes, 1.1-1.4
Axioms of Probability	2.1-2.5
Independence and Conditional Probability	2.6-2.9
Random Variables	3.1-3.4
Multivariate Distributions	3.5-3.8
Expectation of Random Variables	4.1-4.4
MGF and Conditional Expectation	4.5-4.9
Important Discrete Distributions	5.1-5.10
Important Continuous Distributions	6.1-6.8
Functions of Random Variables	7.1-7.6

Grading Policy

Course grade will be based on three exams and homework:
Exam 1, 15%
Exam 2, 25%
Exam 3, 35%
Homework, 25%. Note: 1/2 of the homework grade will be based on whether or not you make a serious attempt at the problems, and 1/2 will be based on how much you get correct.
Note: the complete syllabus is available here:
`http://www.utdallas.edu/~ammann/stat5351_syllabus.pdf`

Math Background

Three important series and an integral

There are three series that are used extensively in this course: binomial, geometric, and exponential power series. The binomial series is defined by,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad 0 \le k \le n.$$

We can evaluate several related series by manipulating the series into the form of a binomial series. For example,

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n}{k} k a^k b^{n-k} &= \sum_{k=1}^{n} \binom{n}{k} k a^k b^{n-k} \\
&= \sum_{k=1}^{n} \frac{kn!}{k!(n-k)!} a^k b^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} a^k b^{n-k} && \text{since } k! = k(k-1)! \\
&= na \sum_{k=1}^{n} \binom{n-1}{k-1} a^{k-1} b^{n-k} && \text{factor out } n \text{ from } n!,\ a \text{ from } a^k \\
&= na \sum_{j=0}^{n-1} \binom{n-1}{j} a^j b^{n-1-j} && (\text{change variables } j = k-1) \\
& \qquad \text{(this is now a binomial series)} \\
&= na(a+b)^{n-1}.
\end{aligned}$$

Another example:

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n}{k} e^{tk} a^k b^{n-k} &= \sum_{k=0}^{n} \binom{n}{k} (ae^t)^k b^{n-k} \\
& \qquad \text{(this is now a binomial series)} \\
&= (ae^t + b)^n.
\end{aligned}$$

The geometric series is defined by

$$\sum_{n=k}^{\infty} r^n = r^k (1-r)^{-1},\ |r| < 1.$$

Similar series can be evaluated by manipulating the series into this form. For example,

$$\begin{aligned}
\sum_{n=0}^{\infty} nr^n &= \sum_{n=1}^{\infty} nr^n \\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{n} r^n \qquad \text{(replace n with } \sum_{\text{k}=1}^{\text{n}} 1) \\
&= \sum\sum_{1\le k\le n<\infty} r^n \\
&= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} r^n \qquad (\textit{interchange order of summations}) \\
&\qquad\qquad \text{(the inner sum is now a geometric series)} \\
&= \sum_{k=1}^{\infty} r^k(1-r)^{-1} \\
&= r(1-r)^{-2}.
\end{aligned}$$

The exponential power series is defined by

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

An example of a related series:

$$\begin{aligned}
\sum_{n=0}^{\infty} n\frac{x^n}{n!} &= \sum_{n=1}^{\infty} n\frac{x^n}{n!} \\
&= \sum_{n=1}^{\infty} \frac{x^n}{(n-1)!} \\
&= x\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \qquad \text{(factor out x from x}^{\text{n}}) \\
&= x\sum_{m=0}^{\infty} \frac{x^m}{m!} \qquad \text{(change variables m = n − 1)} \\
&= xe^x.
\end{aligned}$$

The *Gamma* function is a very important function in mathematics and probability. It is defined by the *Gamma* integral:

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

where $\alpha > 0$. The *Gamma* function satisfies a recursive relationship, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. This relationship can be derived by applying integration by parts to the integral definition of $\Gamma(\alpha+1)$,

$$\begin{aligned}\Gamma(\alpha+1) &= \int_0^\infty x^\alpha e^{-x}dx \\ &\qquad\qquad \text{(integrate by parts with } U = x^\alpha,\ dV = e^{-x}dx) \\ &\qquad\qquad \text{(then } dU = \alpha x^{\alpha-1},\ V = -e^{-x}) \\ &= -x^\alpha e^{-x}|_0^\infty + \alpha\int_0^\infty x^{\alpha-1}e^{-x}dx \\ &= 0 + \alpha\Gamma(\alpha) \\ &= \alpha\Gamma(\alpha).\end{aligned}$$

Note that when α is an integer, then we have the recursion, $\Gamma(n+1) = n\Gamma(n)$. Since

$$\Gamma(1) = \int_0^\infty e^{-x}dx = -e^{-x}|_0^\infty = 1,$$

then we have $\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n!$.

Some related integrals:

$$\begin{aligned}\int_0^\infty x^{\alpha-1}e^{-x/\beta}dx &= \int_0^\infty (y\beta)^{\alpha-1}e^{-y}\beta dy \\ &\qquad\qquad \text{(substitute } y = x/\beta) \\ &= \beta^\alpha\int_0^\infty y^{\alpha-1}e^{-y}dy \\ &= \beta^\alpha\Gamma(\alpha).\end{aligned}$$

$$\begin{aligned}\int_0^\infty e^{-tx}x^{\alpha-1}e^{-x}dx &= \int_0^\infty x^{\alpha-1}e^{-x(1+t)}dx \\ &\qquad\qquad \text{(use the previous example with } \beta = 1/(1+t)) \\ &= (1+t)^{-\alpha}\Gamma(\alpha).\end{aligned}$$

Study Problems

Evaluate the following:

1.

$$\sum_{k=0}^{n}\binom{n}{k}k(k-1)a^k b^{n-k},$$

where $n \geq 2$.

2.

$$\sum_{n=0}^{\infty} n^2 \frac{x^n}{n!}$$

Hint: write n^2 as $n(n-1)+n$.

3.

$$\int_0^{\infty} x^r e^{-\theta x} x^{\alpha-1} e^{-x} dx.$$

Counting methods

As noted previously, when working with experiments that have equally likely outcomes, it is only necessary to count the number of outcomes contained in events to determine their probabilities. Events in many such experiments involve the selection of objects from a population. There are different methods used for counting outcomes for such situations depending on whether or not the selection order of the selected objects is recognized and whether a selected objected is returned (**selection with replacement**) to the population before the next selection is made or not returned (**selection without replacement**). We use the term **permutation** to refer to selection of objects in which selection order is distinguished and use the term **combinations** to refer to the case in which selection order is not distinguished. We will consider here three of these methods, permutations with and without replacement, and combinations without replacement.

Permutations without replacement

If the object selected is not returned to the population before the next object is selected, then an object can appear in the selected subset no more than once. There are n choices in the population to fill the first position, but then that leaves *n-1* choices in the population to fill the second position. Therefore, there are $n(n-1)$ ways to fill the first two positions. Continuing this argument, we can see that there are $n(n-1)\cdots(n+1-k)$ ways to select k objects without replacement from a population of n objects when selection order is distinguished. This number is commonly expressed using factorial notation,

$$n(n-1)\cdots(n+1-k) = \frac{n!}{(n-k)!}$$

Permutations with replacement

This case occurs when we wish to select k objects with replacement from a population of n objects and selection order is distinguished. Replacement implies that the same object could be selected multiple times. What is required is to count the number of distinct sets of

k objects could be selected in this way. We can view this selection process by considering the ways in which each of the positions, $1, \cdots, n$, of the set are filled. Note that there are n choices in the population to fill the first position, and since the object selected for this first position is then returned to the population, there are n choices available for the second selection as well. Therefore, there are n^2 ways to fill the first two positions. Continuing this argument, we can see that there are n^k ways to select k objects with replacement from a population of n distinguishable objects.

Combinations without replacement

The only difference between this case and the case of permutations without replacement is that the selection order of the k selected objects is not distinguished here. This implies that a different arrangement of the same objects is not counted for this case and so this case involves simply selecting a subset of size k from the population. Therefore, we can view the number of permutations without replacement as a two-stage process: first select a subset (combinations without replacement) and then generate every possible rearrangement of each of these subsets. Note that the number of ways to generate every possible rearrangement of k objects is equivalent to counting the number of permutations without replacement of k objects selected from a population of size k, and so is equal to *k!*. Denote by *C(n,k)* the number of combinations without replacement. Then we have,

$$\frac{n!}{(n-k)!} = C(n,k)k!.$$

Hence,

$$C(n,k) = \frac{n!}{k!(n-k)!}.$$

This quantity is usually denoted by

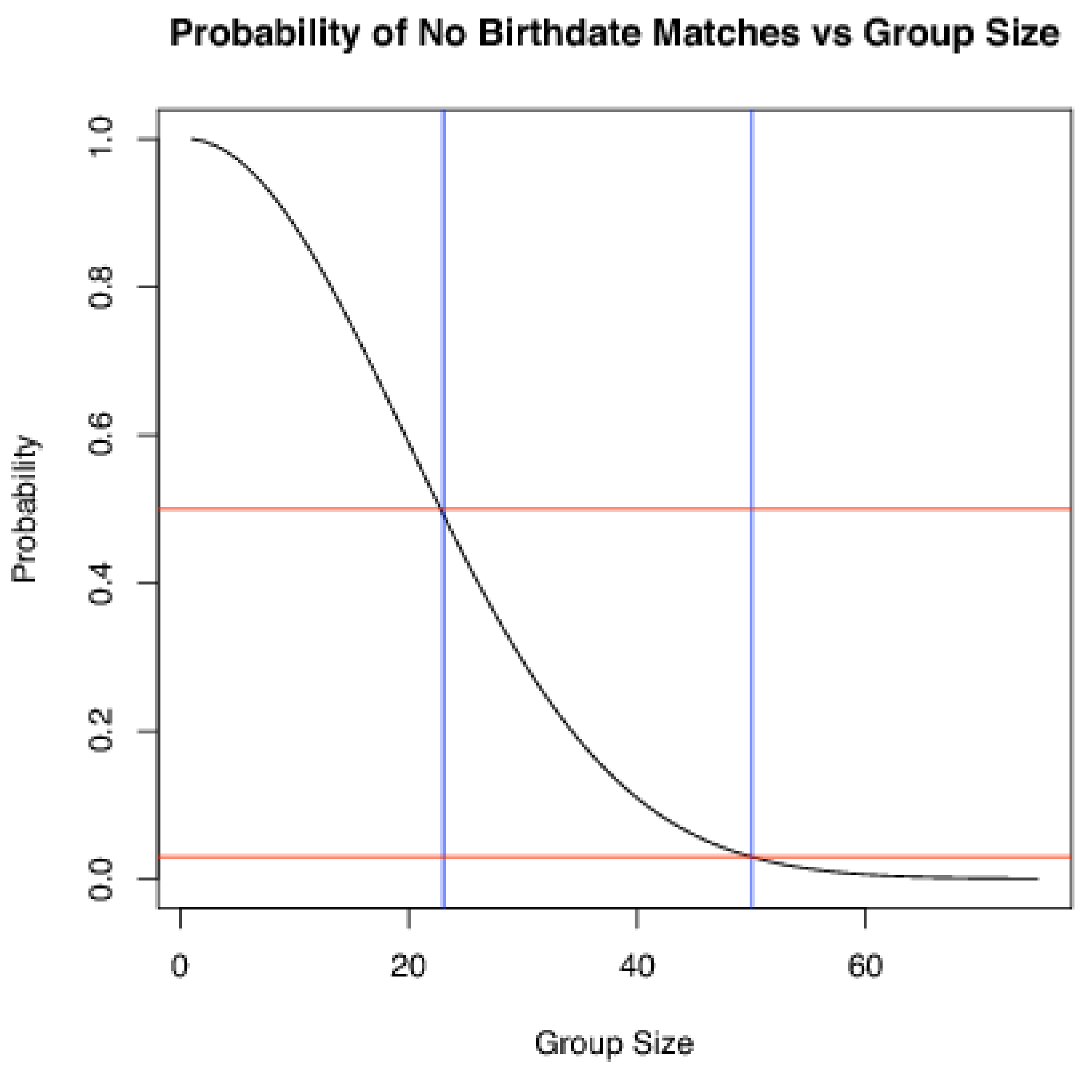
$$\binom{n}{k}$$

Examples

Birthday Problem. This is a classic example of how probability in some applications does not coincide with our intuition. Suppose we have a class of n individuals and want to determine the probability that there is at least one pair of individuals who have the same birthday. To simplify this problem, we will ignore birthdays that occurred on Feb. 29 during a leap year and count those as occurring on March 1. The model we will assume for this problem treats an individual's birthdate as if it was randomly selected from the set of 365 possible birthdays. Therefore, the experiment in which each individual selects a birthdate is an experiment with equally likely outcomes. Therefore, we must count the number of ways to select n birthdays from the population of 365 possible birthdays, and then count the number of ways to select n birthdays with at least one matching pair. It turns out to be easier to count the number of ways to select n birthdays with no matches. This is equivalent to counting the number of permutations without replacement of k objects selected from a population of 365 objects. This number is therefore $365!/(365-n)!$. The total number of possible birthdates for this group is equivalent to the number of permutations with replacement of n birthdates from the population of 365 possible birthdates. This gives,

$$P(\text{no birthdate matches}) = \frac{365!}{(365-n)!365^n}$$

A plot of this probability as a function of n is given below. Note that when n=23, there is about a 50% probability of no matches in the group, and when n=50, there is about a 3% chance of no matches in the group.


Probability of No Birthdate Matches vs Group Size
Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
Group Size

Lottery Games. Texas has a state lottery game which was originally configured so that to win the grand prize, you would need to match six numbers selected from the integers $1, 2, \cdots, 50$ in any order. The experiment in which the six numbered balls are selected can be modelled as an experiment with equally likely outcomes. Therefore, the probability that a particular selection of six numbers will win can be obtained by counting the number of subsets of size six that can be selected from the population consisting of the first 50 integers. That number is,

$$\begin{aligned} \binom{50}{6} &= \frac{50 \cdot 49 \cdot 48 \cdot 47 \cdot 46 \cdot 45}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} \\ &= 15890700. \end{aligned}$$

So the chance of winning this game was *1/15890700 = 0.00000006*. The rules have been changed so that now the six numbers are selected from the first 54 integers. The number of subsets of size six that can be selected from this population has increased to

$$\begin{aligned} \binom{54}{6} &= \frac{54 \cdot 53 \cdot 52 \cdot 51 \cdot 50 \cdot 49}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} \\ &= 25827165, \end{aligned}$$

so the probability of winning the grand prize has been reduced to *1/25827165 = 0.00000004*.

A ticket that matches exactly three of the numbers wins $3. What was the impact on the probability of winning $3 by this change? To answer this question, we must count the number of outcomes that match exactly three numbers. Note that such a match is obtained by selecting a subset of size 3 from the six winning numbers and then selecting a subset of size 3 from the remaining numbers. Under the original rules this number is

$$\begin{aligned} \binom{6}{3}\binom{44}{3} &= \frac{6 \cdot 5 \cdot 4 \cdot 44 \cdot 43 \cdot 42}{3 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} \\ &= 264880. \end{aligned}$$

Under the new rules, this number is

$$\begin{aligned} \binom{6}{3}\binom{48}{3} &= \frac{6 \cdot 5 \cdot 4 \cdot 48 \cdot 47 \cdot 46}{3 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} \\ &= 345920. \end{aligned}$$

The probability of winning this prize is then *264880/15890700 = 0.0167* under the old rules and is *345920/25827165 = 0.0134* under the new rules. Hence, the probability of winning this prize has been reduced by 20%.

Subpopulation selection. Suppose a committee consists of 40 males and 20 females and must select a subcommittee of 5 members. It decides to make this selection randomly. What is the probability that all 5 members of the subcommittee will be female? What is the probability that at least 2 member of the subcommittee will be male? First note that an outcome of this experiment is a set of 5 members selected without replacement from the committee, and this experiment has equally likely outcomes. To answer the questions, we will first obtain the probability that exactly k members of the subcommittee will be female for $0 \leq k \leq 5$. Note that if k members are female, then *5-k* members will be male. Hence, the number of outcomes contained in the event that exactly k members are female can be obtained by counting the number of ways to select a subset of size k from the 20 females and multiplying that times the number of ways to select a subset of size *5-k* from the 40 males. Since order of selection does not count, this number is then

$$\binom{20}{k}\binom{40}{5-k}.$$

The number of outcomes in the sample space is the total number of ways to select a subset of size 5 from the 60 committee members, so the probability that exactly k are female is,

$$P(k) = \frac{\binom{20}{k}\binom{40}{5-k}}{\binom{60}{5}}.$$

We can now answer the questions.

$$\begin{aligned} P(5\ females) &= P(5) = \frac{\binom{20}{5}\binom{40}{0}}{\binom{60}{5}} \\ &= \frac{20!5!55!}{5!15!60!} \\ &= \frac{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} \\ &= 0.0028. \end{aligned}$$

$$\begin{aligned} P(at\ least\ 2\ males) &= P(no\ more\ than\ 3\ females) = P(0) + P(1) + P(2) + P(3) \\ &= 1 - P(4) - P(5). \end{aligned}$$

$$\begin{aligned} P(4) &= \frac{\binom{20}{4}\binom{40}{1}}{\binom{60}{5}} \\ &= \frac{20 \cdot 19 \cdot 18 \cdot 17 \cdot 40 \cdot 5}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} \\ &= 0.0355. \end{aligned}$$

So, $P(at\ least\ 2\ males) = 1 - 0.0355 - 0.0028 = 1 - 0.0383 = .9617$.

Binomial coefficients, revisited

The counting methods described above can be applied directly to selection and/or ordering of subsets taken from populations with distinct items. The previous example of subpopulation selection is an example of selection from a population of items that are not all distinct.

Suppose we have a population that contains r distinct types, with n_i individuals of type i, $1 \le i \le r$. Then the population size is,

$$n = \sum_{i=1}^{r} n_i.$$

The number of distinct subsets selected without replacement from this population that contain k_i individuals of type i, where $0 \le k_i \le n_i$, is given by

$$\binom{n_1}{k_1} \cdots \binom{n_r}{k_r}.$$

Consider the special case in which $r = 2$. The number of subsets of size m selected without replacement from this population, if we ignore the types of those selected, is the binomial coefficient,

$$\binom{n}{m}.$$

We can select such a sample in a two-step process. We will assume that $m \le n_i$, $i = 1, 2$. For each $0 \le k \le m$, select k type 1 items and m-k type 2 items. The number of subsets that contain k type 1 and m-k type 2 items is

$$\binom{n_1}{k}\binom{n_2}{m-k}.$$

Let A denote the collection of all subsets of size m selected without replacement, and let A_k denote the collection of all subsets that contain k type 1 and m-k type 2 items. Then these sets are disjoint and

$$\bigcup_{k=0}^{m} A_k = A.$$

Hence,

$$\binom{n}{m} = \sum_{k=0}^{m} \binom{n_1}{k}\binom{n_2}{m-k}.$$

Note that this is Theorem 1.12 in the text.

Now suppose we wish to order a population with r distinct types. A specific ordering of this population can be obtained by specifying the positions for each type among the set of all positions, $1, \ldots, n$. The number of ways this can be done corresponds to the number of ways we can partition this set into subsets of size n_i, $1 \le i \le r$. Note that when the positions for types $1, \ldots, r-1$ are specified, then the remaining positions would be for type

r. We will obtain these counts sequentially. The number of ways we can select n_1 postions from the set of all n positions is,

$$\binom{n}{n_1}.$$

Now we must select n_2 postions for type 2 items from the remaining $n - n_1$ positions. The number of such positions is then,

$$\binom{n-n_1}{n_2}.$$

The total number of way to select positions for type 1 and type 2 items is then,

$$\begin{aligned}\binom{n}{n_1}\binom{n-n_1}{n_2} &= \left[\frac{n!}{n_1!(n-n_1)!}\right]\left[\frac{(n-n_1)!}{n_2!(n-n_1-n_2)!}\right] \\ &= \frac{n!}{n_1!n_2!(n-n_1-n_2)!}.\end{aligned}$$

Continuing this process, the number of ways to order this population is then

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

Example. How many distinct arrangements of the letters in the word *statistics* are there? Note that the letter s occurs 3 times, t occurs 3 times, a occurs 1 time, i occurs 2 times, c occurs 1 time for a total of 10 characters. Therefore, the number of distinct arrangements of these letters is,

$$\frac{10!}{3!3!1!2!1!} = 50400.$$

Homework Assignments

Homework 1

1. Exercise 1.30, p. 19
2. Exercise 1.36, p. 20
3. Exercise 2.46, p. 58
4. Exercise 2.62, p. 60
5. Exercise 2.90, p. 64

6. Examination of recent employment records for a company revealed that during the 4th quarter of 2000, the company hired 10 entry-level management employees. One of the 10 hired was female. There were 80 qualified applicants for these positions, 36 of which were female. Treat the selection of the 10 who were hired as the result of a random experiment in which a sample of size 10 was randomly selected without replacement from the population of 80 qualified applicants. This assumption corresponds to the assumtion that the applicants are equally qualified. In particular, it means that the qualifications for the positions were not related to the gender of the applicant. Determine the probability of hiring no more than 1 female under this assumption. What would you say about the company based on this probability?

7. From a batch of 20 components a sample of 6 is randomly selected without replacement for inspection. If there are 5 defective components in the batch, what is the probability that a sample contains

   (a) exactly 1 defective?

   (b) at least 1 defective?

   (c) all five defectives?

Solutions for Homework 1

1. Exercise 1.30, p. 19.
   (a) 6.
   (b) Since we don't care about order, we just need to select a point for 2 of the dice and a point from the remaining points. This gives
   $$(6)(5) = 30.$$
   (c) Select 3 different points,
   $$\binom{6}{3} = 20.$$
   (d) 6+30+20=56.

2. Exercise 1.36, p. 20.
   (a)
   $$\binom{14}{2}\binom{1}{1} = 91.$$
   (b)
   $$\binom{14}{3} = 364.$$

3. Exercise 2.46, p. 58
   1: Driver has liability and collision insurance
   2: Driver has liability but not collision insurance
   3: Driver has collision but not liability insurance
   4. Driver has no liability and no liability insurance.

4. Exercise 2.62, p. 60
   (a) First select the two denominations for the pairs, then select the suits for each pair, then select one card from remaining 11 denominations:
   $$\binom{13}{2}\binom{4}{2}\binom{4}{2}(44) = 123552.$$
   Now divide by number of 5-card hands to get probability:
   $$P(2\ pair) = \frac{\binom{13}{2}\binom{4}{2}\binom{4}{2}(44)}{\binom{52}{5}} = 0.0475.$$

(b) First select the denomination for the four of a kind and then select one card from remaining:

$$\binom{13}{1}(48) = 624.$$

This gives

$$P(4\ of\ a\ kind) = \frac{\binom{13}{1}(48)}{\binom{52}{5}} = 0.00024.$$

5. Exercise 2.90, p. 64. This is an experiment with equally likely outcomes, so we need to count the number of ways to select 3 eggs out of the 110 good eggs in the crate and then divide this by the number of ways to select 3 eggs out of all 120 eggs in the crate. This gives,

$$\begin{aligned} P(\text{crate is shipped}) &= \frac{\binom{110}{3}}{\binom{120}{3}} \\ &= \frac{110 * 109 * 108}{120 * 119 * 118} = 0.768. \end{aligned}$$

6. Let N denote the number of females hired. Then

$$P(N \leq 1) = P(N = 0) + P(N = 1).$$

Now,

$$\begin{aligned} P(N = 0) &= P(\text{10 males were hired}) \\ &= \frac{\binom{44}{10}}{\binom{80}{10}} \\ &= 0.0015, \end{aligned}$$

$$\begin{aligned} P(N = 1) &= P(\text{9 males were hired and 1 female was hired}) \\ &= \frac{\binom{44}{9}\binom{36}{1}}{\binom{80}{10}} \\ &= 0.0155. \end{aligned}$$

So, $P(N \leq 1) = 0.0015 + 0.0155 = 0.017$. It is highly unlikely (p=0.017) this result could have occurred by chance alone, and so this provides some evidence for bias in their hiring selection process.

7. Let N denote number of defective components selected.

(a)

$$P(N = 1) = \frac{\binom{5}{1}\binom{15}{5}}{\binom{20}{6}} = 0.387.$$

(b)

$$P(N \geq 1) = 1 - P(N = 0) = 1 - \frac{\binom{15}{6}}{\binom{20}{6}} = 1 - 0.129 = 0.871.$$

(c)

$$P(N = 5) = \frac{\binom{5}{5}\binom{15}{1}}{\binom{20}{6}} = 0.000387.$$

Homework 2

Due date: 9/22/2010

1. Exercise 2.21, p. 52
2. Exercise 2.22, p. 52
3. Exercise 2.106, p. 66
4. Determine whether or not each of the following statements satisfies the laws of probability. For those that do not, state how the laws are violated.
   (a) $P(A) = 0.9$, $P(A|B) = 0.5$,$P(B|A) = 0.6$.
   (b) A and B are independent, $P(A \cup B) = 0.8$, $P(A) = 0.6$.
   (c) $P(A^c|B) = 0.3$, $P(A|B^c) = 0.7$, $P(B) = 0.9$, $P(A) = 0.6$.
5. A company has been running a television advertisement for one of its new products. A survey was conducted, and based on its results, it was concluded that an individual buys the product with probability 56% if he/she saw the advertisement, and buys with probability 8% if he/she did not see it. Twenty-five percent of people saw the advertisement.
   (a) What is the probability that a randomly selected individual will buy the new product?
   (b) What is the probability that at least one of five randomly selected individuals will buy the new product? Assume the individuals are acting independently.
6. Exercise 3.8, p. 81
7. Exercise 3.12, p. 81
8. Exercise 3.24, p. 90
9. Exercise 3.26, p. 90
10. Exercise 3.84, p. 121
11. Exercise 3.92, p. 121
12. Exercise 3.96, p. 122

Solutions for Homework 2

1. Exercise 2.21, p. 52.

$$\begin{aligned} P(A|B) &= \frac{P(A\cap B)}{P(B)} \quad \text{(from the def. of cond. prob.)} \\ &= \frac{P(B|A)P(A)}{P(B)} \quad \text{(mult. rule)} \\ &= \frac{P(B)P(A)}{P(B)} \\ &= P(A). \end{aligned}$$

2. Exercise 2.22, p. 52. Since A and B are independent, then $P(A\cap B) = P(A)P(B)$.
(a). From the Law of Total Probability,

$$\begin{aligned} P(B) &= P(A\cap B) + P(A'\cap B) \\ &= P(A)P(B) + P(A'\cap B). \end{aligned}$$

Hence,

$$\begin{aligned} P(A'\cap B) &= P(B) - P(A)P(B) \\ &= [1 - P(A)]P(B) \\ &= P(A')P(B). \end{aligned}$$

This implies that A' and B are independent.
(b). This proof is similar. From part **(a)**,

$$\begin{aligned} P(A') &= P(A'\cap B) + P(A'\cap B') \\ &= P(A')P(B) + P(A'\cap B'), \end{aligned}$$

and so,

$$\begin{aligned} P(A'\cap B') &= P(A') - P(A')P(B) \\ &= P(A')[1 - P(B)] \\ &= P(A')P(B'). \end{aligned}$$

3. Exercise 2.106, p. 66. First translate. Here is another statement of this problem in English. 4 percent of accidents are due to faulty brakes. The reports from 82 percent of accidents caused by faulty brakes correctly list the cause as faulty brakes. The reports from 3 percent of accidents not caused by faulty brakes incorrectly list the cause as

faulty brakes. Note that the 4 percent refers to all accidents, and so that corresponds to an ordinary probability,

$$P(\text{brakes caused accident}) = 0.04.$$

However, the 82 percent only refers to accidents caused by faulty brakes, and so that corresponds to a conditional probability,

$$P(\text{cause listed as faulty brakes}|\text{brakes caused accident}) = 0.82.$$

Likewise, the 3 percent refers only to accidents not caused by faulty brakes, and so it also corresponds to a conditional probability,

$$P(\text{cause listed as faulty brakes}|\text{brakes did not cause accident}) = 0.03.$$

Let

$$A = \{\text{brakes caused accident}\},$$

$$B = \{\text{cause listed as faulty brakes}\}.$$

Then we have,

$$P(A) = 0.04, \quad P(B|A) = 0.82, \quad P(B|A') = 0.03.$$

(a). We are asked to find $P(B)$,

$$\begin{aligned} P(B) &= P(B \cap A) + P(B \cap A') \\ &= P(B|A)P(A) + P(B|A')P(A') \\ &= (0.82)(0.04) + (.03)(.96) \\ &= 0.0616. \end{aligned}$$

(b). This question translates to: find $P(A|B)$.

$$\begin{aligned} P(A|B) &= \frac{P(B \cap A)}{P(B)} \\ &= \frac{(0.82)(0.04)}{0.0616} \\ &= 0.532. \end{aligned}$$

4. **(a)**. Since $P(A) = 0.9$ and $P(B|A) = 0.6$, then $P(A \cap B) = (.9)(.6) = 0.54$. But this would imply that

$$0.54 = P(A \cap B) = P(A|B)P(B) = 0.5P(B),$$

and so

$$P(B) = \frac{.54}{.5} = 1.08,$$

which violates the laws of probability.
(b). Since A and B are independent, then $P(A \cap B) = P(A)P(B)$. Hence,

$$0.8 = P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = 0.6 + P(B) - .6P(B).$$

This gives,

$$0.2 = .4P(B),$$

which implies that $P(B) = 0.5$. Laws of probability are satisfied.
(c). First note that

$$P(A^c \cap B) = P(A^c|B)P(B) = (0.3)(0.9) = 0.27,$$

and

$$P(A \cap B^c) = P(A|B^c)P(B^c) = (0.7)(0.1) = 0.07.$$

But

$$P(A \cap B) = P(B) - P(A^c \cap B) = 0.9 - 0.27 = 0.63 > 0.6 = P(A),$$

which violates laws of probability.

5. First translate:
an individual buys the product with probability 56% if he/she saw the advertisement: since a condition is imposed on the individuals who buy, this is a conditional probability, $P(buys|saw\ ad) = 0.56$.
buys with probability 8% if he/she did not see it: since a condition is imposed on the individuals who buy, this is a conditional probability, $P(buys|not\ see\ ad) = 0.08$.
Twenty-five percent of people saw the advertisement: no conditions, so this is an ordinary probability, $P(saw\ ad) = 0.25$.

   (a) Translate: no conditions so this question is an ordinary probability, $P(buy)$. Use Theorem of Total Probability and multiplicative law:

$$\begin{aligned} P(buy) &= P(buys|saw\ ad)P(saw\ ad) + P(buys|not\ see\ ad\ ad)P(not\ see\ ad) \\ &= (0.56)(0.25) + (0.08)(0.75) \\ &= 0.20. \end{aligned}$$

(b) Assumption of independence implies that

$$\begin{aligned} P(\textit{at least } 1 \textit{ buys}) &= 1 - P(\textit{none buy}) \\ &= 1 - P(\textit{not buy})^5 \\ &= 1 - .8^5 \\ &= 0.672. \end{aligned}$$

6. Exercise 3.8, p. 81
   1.

$$F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{n \to \infty} P(X \le -n).$$

Let $A_n = \{X \le -n\},\ n \ge 1$. Then $A_{n+1} \subset A_n$, $n \ge 1$, and

$$\bigcap_{n=1}^{\infty} A_n = \lim_{n \to \infty} A_n = \emptyset.$$

Hence, by the continuity theorem,

$$F(-\infty) = \lim_{n \to \infty} P(A_n) = P(\lim_{n \to \infty} A_n) = P(\emptyset) = 0.$$

Similarly, $B_n = \{X \le n\},\ n \ge 1$, defines an increasing sequence of events with

$$\bigcup_{n=1}^{\infty} B_n = \lim_{n \to \infty} B_n = \Omega,$$

and so

$$F(\infty) = \lim_{x \to \infty} F(x) = \lim_{n \to \infty} P(B_n) = P(\lim_{n \to \infty} B_n) = P(\Omega) = 1.$$

2. If $a < b$, then $X \le a \subset X \le b$, and so by the monotonicity theorem,

$$F(a) = P(X \le a) \le P(X \le b) = F(b).$$

7. Exercise 3.12, p. 81
   This is the d.f. of a discrete r.v. with

$$\begin{aligned} P(X = 1) &= \frac{1}{3}; \quad P(X = 4) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6} \\ P(X = 6) &= \frac{5}{6} - \frac{1}{2} = \frac{1}{3}; \quad P(X = 10) = 1 - \frac{5}{6} = \frac{1}{6} \end{aligned}$$

(a)

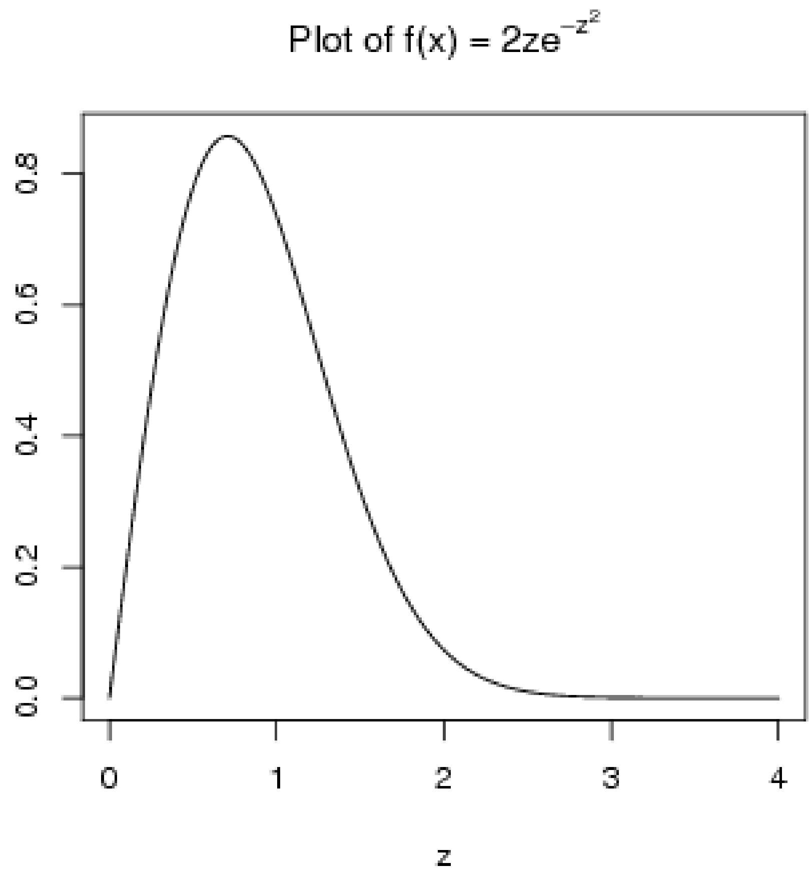
$$P(2 < X \le 6) = P(X = 4) + P(X = 6) = \frac{1}{2}.$$

(b) $P(X = 4) = 1/6$.
(c) See above.

8. Exercise 3.24, p. 90

$$\int_0^\infty ze^{-z^2}dz = -\frac{1}{2}e^{-z^2}\Big|_0^\infty = \frac{1}{2}.$$

Therefore, $k = 2$.



9. Exercise 3.26, p. 90

$$P(X < \frac{1}{4}) = \int_0^{1/4} 6x(1-x)dx = (3x^2 - 2x^3)\Big|_0^{1/4} = \frac{5}{32}.$$

$$P(X > \frac{1}{2}) = \int_{1/2}^{1} 6x(1-x)dx = (3x^2 - 2x^3)\Big|_{1/2}^{1} = \frac{1}{2}.$$

10. Exercise 3.84, p. 121
The outcomes of this experiment are: $\{12, 13, 14, 23, 24, 34\}$, and each of these outcomes has probability $1/6$. The sum of the numbers on the balls drawn corresponding to these

outcomes is, respectively, $\{3, 4, 5, 5, 6, 7\}$.

(a) The p.m.f. for the sum of the numbers on the balls drawn is,

$$P(X = 3) = 1/6,\ \ P(X = 4) = 1/6,\ \ P(X = 5) = 1/3,\ \ P(X = 6) = 1/6,\ \ P(X = 7) = 1/6,$$

and $P(X = x) = 0$ for any other value of x.

(b) The d.f. is

$$F(x) = \begin{cases} 0, & x < 3 \\ \frac{1}{6}, & 3 \leq x < 4 \\ \frac{1}{3}, & 4 \leq x < 5 \\ \frac{2}{3}, & 5 \leq x < 6 \\ \frac{5}{6}, & 6 \leq x < 7 \\ 1, & x \geq 7. \end{cases}$$

11. Exercise 3.92, p. 121

(a)

$$\begin{aligned} P(X \leq -2) &= \int_{-6}^{-2} \frac{1}{288}(36 - x^2)dx \\ &= \frac{1}{288}(36x - \frac{1}{3}x^3)\Big|_{-6}^{-2} \\ &= \frac{7}{27}. \end{aligned}$$

(b)

$$\begin{aligned} P(X \geq 1) &= \int_{1}^{6} \frac{1}{288}(36 - x^2)dx \\ &= \frac{1}{288}(36x - \frac{1}{3}x^3)\Big|_{1}^{6} \\ &= \frac{325}{864} \approx 0.376. \end{aligned}$$

(c)

$$\begin{aligned} P(-3 \leq X \leq -1) &= \int_{-3}^{-1} \frac{1}{288}(36 - x^2)dx \\ &= \frac{1}{288}(36x - \frac{1}{3}x^3)\Big|_{-3}^{-1} \\ &= \frac{190}{864} \approx 0.220. \end{aligned}$$

(d) 0.

12. Exercise 3.96, p. 122

(a) $P(X > 10) = 1 - P(X \leq 10) = 1 - F(10) = 1 - (1 - .25) = 0.25.$

(b) $P(X < 8) = F(8) = 1 - 25/64 = 39/64.$

(c) $P(12 \leq X \leq 15) = F(15) - F(12) = 25/144 - 25/225 = 1/16.$

Homework 3

Due date: Oct. 25, 2010.

Show your work to get full credit for these problems.

1. Exercise 3.49, p. 102
2. Exercise 3.50, p. 102
3. Exercise 3.51, p. 102
4. Exercise 3.52, p. 102
5. Exercise 3.53, p. 102
6. Exercise 3.58, p. 103
7. Exercise 3.102, p. 123
8. (*Extra credit*) Let X and Y be independent random variables that have the same probability mass function given by

   $$f(k) = (1-p)p^k, \; k \geq 0,$$

   where p is a constant with $0 < p < 1$. Let $U = \min(X, Y)$, $V = \max(X, Y)$, and $W = V - U$.

   (a) Find the joint probability mass function of U and W.

   (b) Show that U and W are independent.

   Hint: first obtain $P(U = j, V = k)$ and note that $V \geq U$. Also note that the case in which $k = j$ must be handled differently than the case in which $k > j$.

9. Exercise 4.8, p. 137
10. Exercise 4.46, p. 157
11. Exercise 4.52, p. 158
12. Exercise 4.60, p. 161
13. A fund-raising event for a charitable organization includes a game in which contestants throw darts at a target. The target consists of an outer ring, an inner ring, and a circle inside the inner ring. Suppose the organization has several people throw darts at this target to get approximate values for the probabilities of hitting these areas and finds that the outer ring is hit approximately 30% of the time, the inner ring is hit approximately 10% of the time, and the circle is hit approximately 5% of the time. The rest of the time the target is missed completely. They plan to pay \$0.50 if a dart hits the outer ring, \$1.00 if it hits the inner ring, and \$5.00 if it hits the circle. If a dart misses the target, then the contestant does not win anything.

(a) If the organization charges \$1 to throw a dart, what is their expected net profit per dart based on these probabilities?

(b) At least 75% of the time this organization can expect their net profit per dart to be within what interval?

14. An insurance company sells a policy that covers laptops sold by the University's technology store. Suppose this policy pays \$100 for minor damage other than video screen, \$500 for major damage other than video screen, \$800 for video screen damage, \$1200 if the laptop must be replaced, and these are the only claims under this policy. Based on past experience, the insurance company has found that the probabilities for these events are 0.20 for minor damage, 0.10 for major damage, 0.07 for video damage, and 0.05 for replacement.

    (a) What proportion of policies receive no payments?

    (b) What premium should the insurance company charge to have an expected net gain of \$100 per policy?

    (c) What is the s.d. of the insurance company's net gain?

Solutions for Homework 3

1. Exercise 3.49, p. 102

$$\begin{aligned}
\int_0^1 \int_{-x}^{x} x(x-y)dydx &= \int_0^1 \left[x(xy - \frac{1}{2}y^2\Big|_{-x}^{x} \right] dx \\
&= \int_0^1 x(x^2 - \frac{1}{2}x^2 + x^2 + \frac{1}{2}x^2)dx \\
&= \int_0^1 2x^3 dx \\
&= \frac{1}{2}x^4\Big|_0^1 \\
&= \frac{1}{2}.
\end{aligned}$$

Therefore, $k = 2$.

2. Exercise 3.50, p. 102
The region $0 < X + Y < 1/2$ is entirely within the region over which the density is positive, so we just integrate the density over that region.

$$\begin{aligned}
P(X+Y < 1/2) &= \int_0^{.5} \int_0^{y-.5} 24xydxdy \\
&= \int_0^{.5} 12y \left[x^2\Big|_0^{.5-y} \right] dy \\
&= \int_0^{.5} (3y - 12y^2 + 12y^3)dy \\
&= (\frac{3}{2}y^2 - 4y^3 + 3y^4)\Big|_0^{.5} \\
&= \frac{1}{16}.
\end{aligned}$$

3. Exercise 3.51, p. 102

$$\begin{aligned}
P(X \le 1/2, Y \le 1/2) &= \int_0^{1/2} \int_0^{1/2} 2dxdy \\
&= \int_0^{1/2} 2(1/2)dy = \int_0^{1/2} dy \\
&= 1/2.
\end{aligned}$$

To find $P(X+Y>2/3)$ we must integrate the joint density over the region between the lines, $x+y=2/3$ and $x+y=1$. Note that the lower bound for x is $2/3-y$ when $0<y<2/3$, but is 0 when $2/3 \leq y \leq 1$. This means that we must split the region of integration into these two sections.

$$\begin{aligned}
P(X+Y>2/3) &= \int_0^{2/3}\int_{2/3-y}^{1-y} 2dxdy + \int_{2/3}^{1}\int_0^{1-y} 2dxdy \\
&= \int_0^{2/3} 2[(1-y)-(2/3-y)]dy + \int_{2/3}^{1} 2(1-y)dy \\
&= \int_0^{2/3} \frac{2}{3}dy + [2y-y^2]\Big|_{2/3}^{1} \\
&= \frac{4}{9} + (2-1-\frac{4}{3}+\frac{4}{9}) \\
&= \frac{5}{9}.
\end{aligned}$$

To find $P(X>2Y)$ we must integrate the joint density over the region between the lines, $x=2y$ and $x+y=1$. Note that if we integrate with respect to x first, then we don't need to split the region of integration, but if we integrate with respect to y first, then we will need to split the region.

$$\begin{aligned}
P(X>2Y) &= \int_0^{1/3}\int_{2y}^{1-y} 2dxdy \\
&= \int_0^{1/3} 2(1-y-2y)dy \\
&= \int_0^{1/3} 2(1-3y)dy \\
&= (2y-3y^2)\Big|_0^{1/3} \\
&= \frac{1}{3}.
\end{aligned}$$

4. Exercise 3.52, p. 102
Since the point at which we are asked to evaluate the joint d.f. is inside the region over which the density is positive, we can obtain the d.f. directly from the density function

by integrating over the rectangular region, $0 < s < x$, $0 < t < y$.

$$\begin{aligned} F(x,y) &= P(X \le x, Y \le y) \\ &= \int_0^y \int_0^x 2dxdy \\ &= \int_0^y 2xdy \\ &= 2xy. \end{aligned}$$

Note that if (s,t) is outside the region $x > 0, y > 0, x + y < 1$, then the region of integration would not be retangular.

5. Exercise 3.53, p. 102
   To satisfy the requirements that $0 < x < y < 1$ and $x + y > 1/2$, region of integration must be split into two subregions:

   $$1/4 < y < 1/2,\ 1/2 - y < x < y$$

   and

   $$1/2 < y < 1,\ 0 < x < y.$$

   $$\begin{aligned} P(X + Y > \frac{1}{2}) &= \int_{1/4}^{1/2} \int_{1/2-y}^{y} \frac{1}{y} dxdy + \int_{1/2}^{1} \int_0^y \frac{1}{y} dxdy \\ &= \int_{1/4}^{1/2} \frac{2y - 1/2}{y} dy + \int_{1/2}^{1} dy \\ &= \frac{1}{2} + \int_{1/4}^{1/2} (2 - \frac{1}{2y}) dy \\ &= 1 - \frac{1}{2}\log(\frac{1}{2}) + \frac{1}{2}\log(\frac{1}{4}) \\ &= 1 - \frac{1}{2}\log(2). \end{aligned}$$

6. Exercise 3.58, p. 103
   First note that $F(x,y)$ is the probability of the intersection of the events, $X \le x$ and $Y \le y$. Let

   $$A = \{X \le a\},\ B = \{X \le b\},\ C = \{Y \le c\},\ D = \{Y \le d\},$$

Then

$$\begin{aligned} P(a < X \le b, c < Y \le d) &= P(B \bigcap A^c \bigcap D \bigcap C^c) \\ &= P(B \bigcap A^c \bigcap D) - P(B \bigcap A^c \bigcap C) \\ &= P(B \bigcap D) - P(A \bigcap D) - P(B \bigcap C) + P(A \bigcap C) \\ &= F(b,d) - F(a,d) - F(b,c) + F(a,c). \end{aligned}$$

It is helpful to draw these regions.

7. Exercise 3.102, p. 123

$$\begin{aligned} P(P < .3, S > 2) &= \int_{.2}^{.3} \int_{2}^{\infty} 5pe^{-ps} ds dp \\ &= \int_{.2}^{.3} 5e^{-2p} dp \\ &= -2.5e^{-2p} \Big|_{.2}^{.3} \\ &= 2.5(e^{-.4} - e^{-.6}) = 0.304. \end{aligned}$$

$$\begin{aligned} P(.25 < P < .3, S < 1) &= \int_{.25}^{.3} \int_{0}^{1} 5pe^{-ps} ds dp \\ &= \int_{.25}^{.3} 5(1 - e^{-p}) dp \\ &= 0.25 + 5e^{-p} \Big|_{.25}^{.3} \\ &= 0.25 + 5e^{-.3} - 5e^{-.25}) = 0.060. \end{aligned}$$

8. (*Extra credit*) Let X and Y be independent random variables that have the same probability mass function given by

$$f(k) = (1-p)p^k, \ k \ge 0,$$

where p is a constant with $0 < p < 1$. Let $U = \min(X, Y)$, $V = \max(X, Y)$, and $W = V - U$.

   (a) Find the joint probability mass function of U and W.

   (b) Show that U and W are independent.

Hint: first obtain $P(U=j, V=k)$ and note that $V \geq U$. Also note that the case in which $k=j$ must be handled differently than the case in which $k>j$.
Solution. First express the joint p.m.f. of U, W in terms of U, V.

$$P(U=i, W=j) = P(U=i, V-U=j) = P(U=i, V=i+j),$$

for $i \geq 0,\ j \geq 0$. Note that if $j=0$, then

$$\{\min(X,Y) = i = \max(X,Y)\} \iff \{X=i, Y=i\}.$$

However, if $j>0$, then

$$\{\min(X,Y)=i, \max(X,Y)=i+j\} \iff \{X=i, Y=i+j\} \bigcup \{X=i+j, Y=i\}.$$

Therefore,

$$P(U=i, W=0) = P(X=i, Y=i) = (1-p)p^i(1-p)p^i = (1-p)^2p^{2i}, \quad i \geq 0,$$

and for $j>0$,

$$\begin{aligned} P(U=i, W=j) &= P(X=i, Y=i+j) + P(X=i+j, Y=i) \\ &= 2(1-p)^2p^{2i+j}, \quad i \geq 0,\ j \geq 1. \end{aligned}$$

To show independence of U and W, we must obtain their marginal p.m.f.'s.

$$\begin{aligned} P(U=i) &= P(U=i, W=0) + \sum_{j=1}^{\infty} 2(1-p)^2p^{2i+j} \\ &= (1-p)^2p^{2i} + 2(1-p)^2p^{2i}\sum_{j=1}^{\infty} p^j \\ &= (1-p)^2p^{2i} + 2(1-p)^2p^{2i+1}\frac{1}{1-p} \\ &= (1-p)^2p^{2i} + 2(1-p)p^{2i+1} \\ &= (1-p)p^{2i}(1-p+2p) \\ &= (1-p^2)p^{2i},\ i \geq 0. \end{aligned}$$

$$\begin{aligned} P(W=0) &= \sum_{i=0}^{\infty}(1-p)^2p^{2i} \\ &= (1-p)^2\frac{1}{1-p^2} \\ &= (1-p)^2\frac{1}{(1-p)(1+p)} \\ &= \frac{1-p}{1+p}, \end{aligned}$$

and for $j \geq 1$,

$$\begin{aligned}
P(W=j) &= \sum_{i=0}^{\infty} 2(1-p)^2 p^{2i+j} \\
&= 2(1-p)^2 \frac{p^j}{1-p^2} \\
&= 2(1-p)^2 \frac{p^j}{(1-p)(1+p)} \\
&= \frac{2(1-p)p^j}{1+p}, \ \ j \geq 1.
\end{aligned}$$

Since

$$P(U=i)P(W=0) = (1-p^2)p^{2i}\frac{1-p}{1+p} = (1-p)^2 p^{2i} = P(U=i, W=0), \quad i \geq 0,$$

and

$$\begin{aligned}
P(U=i)P(W=j) &= (1-p^2)p^{2i}\frac{2(1-p)p^j}{1+p} \\
&= 2(1-p)^2 p^{2i+j} = P(U=i, W=j), \quad i \geq 0, \ j \geq 1,
\end{aligned}$$

then U, W are independent.

9. Exercise 4.8, p. 137

$$\begin{aligned}
E(X) &= \int_0^1 x^2 dx + \int_1^2 x(2-x)dx \\
&= \frac{1}{3}x^3\Big|_0^1 + (x^2 - \frac{x^3}{3})\Big|_1^2 \\
&= \frac{1}{3} + 4 - \frac{8}{3} - 1 + \frac{1}{3} \\
&= 1.
\end{aligned}$$

10. Exercise 4.46, p. 157
First note that

$$\begin{aligned}
E(U) &= \int_{-1}^0 x(1+x)dx + \int_0^1 x(1-x)dx \\
&= (\frac{x^2}{2} + \frac{x^3}{3})\Big|_{-1}^0 + (\frac{x^2}{2} - \frac{x^3}{3})\Big|_0^1 \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned} E(UV) &= \int_{-1}^{0} x^3(1+x)dx + \int_0^1 x^3(1-x)dx \\ &= (\frac{x^4}{4} + \frac{x^5}{5})\Big|_{-1}^{0} + (\frac{x^4}{4} - \frac{x^5}{5})\Big|_0^1 \\ &= 0. \end{aligned}$$

These imply that

$$\text{Cov}(U, V) = E(UV) - E(U)E(V) = 0.$$

To show that these r.v.'s are not independent, let $A = \{0 \le U < 1/2\}$ and $B = \{0 \le V < 1/4\}$. Then $A \subset B$ and so

$$P(A \bigcap B) = P(A).$$

Since $0 < P(B) < 1$, then

$$P(A \bigcap B) = P(A) < P(A)P(B),$$

and so these r.v.'s are not independent.

11. Exercise 4.52, p. 158
We showed in class that if $Z = a + bX + cY$, where a, b, c are constants, then

$$E(Z) = a + b\mu_X + c\mu_Y,$$

and

$$Var(Z) = b^2\sigma_X^2 + c^2\sigma_Y^2 + 2bc\text{Cov}(X, Y),$$

where $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ are the respective means and variances of X, Y. So

$$Var(X + Y) = \sigma_X^2 + \sigma_Y^2 + 2\text{Cov}(X, Y),$$

$$Var(X - Y) = \sigma_X^2 + \sigma_Y^2 - 2\text{Cov}(X, Y),$$

and

$$\begin{aligned} \text{Cov}(X + Y, X - Y) &= E[(X + Y)(X - Y) - E(X + Y)E(X - Y) \\ &= E(X^2) - E(Y^2) - (EX)^2 + (EY)^2 \\ &= \sigma_X^2 - \sigma_Y^2. \end{aligned}$$

Note that if X and Y have the same variances, then

$$\text{Cov}(X + Y, X - Y) = 0.$$

12. Exercise 4.60, p. 161

$$E(\text{Profit}) = (.15)(3000) + (.35)(1500) + (.35)(0) + (.15)(-1500) = 750.$$

13. A fund-raising event for a charitable organization includes a game in which contestants throw darts at a target. The target consists of an outer ring, an inner ring, and a circle inside the inner ring. Suppose the organization has several people throw darts at this target to get approximate values for the probabilities of hitting these areas and finds that the outer ring is hit approximately 30% of the time, the inner ring is hit approximately 10% of the time, and the circle is hit approximately 5% of the time. The rest of the time the target is missed completely. They plan to pay \$0.50 if a dart hits the outer ring, \$1.00 if it hits the inner ring, and \$5.00 if it hits the circle. If a dart misses the target, then the contestant does not win anything.

    (a) If the organization charges \$1 to throw a dart, what is their expected net profit per dart based on these probabilities?

    (b) At least 75% of the time this organization can expect their net profit per dart to be within what interval?

    Solution. Denote by X the r.v. that represents the payout to a player after throwing a dart. Then,

    $$P(X = 0.50) = 0.30, \quad P(X = 1.00) = 0.10, \quad P(X = 5.00) = 0.05.$$

    Also, these probabilities imply that $P(X = 0) = 0.55$. The r.v. of interest is net profit, so we must express the relationship between payout and net profit. Denote net profit by Y. Since net profit is the price of a dart to throw minus what the organization pays a player, then

    $$Y = 1 - X.$$

    Therefore, $E(Y) = 1 - E(X)$ and $Var(Y) = Var(X)$ since Y is a linear function of X.

    $$E(X) = (.5)(.3) + (1)(.1) + (5)(.05) = 0.50,$$

    $$E(X^2) = (.5^2)(.3) + (1^2)(.1) + (5^2)(.05) = 0.075 + .100 + 1.250 = 1.425$$

    Hence,

    $$Var(Y) = Var(X) = 1.425 - .5^2 = 1.175,$$

    and so,

    $$sd(Y) = \sqrt{1.175} = 1.08.$$

    By Chebychev's Theorem, at least 75% of the time the organization can expect their net profit per dart to be in the interval $0.50 \pm 2(1.08)$, or $(-1.66, 2.66)$.

14. An insurance company sells a policy that covers laptops sold by the University's technology store. Suppose this policy pays \$100 for minor damage other than video screen, \$500 for major damage other than video screen, \$800 for video screen damage, \$1200 if the laptop must be replaced, and these are the only claims under this policy. Based on past experience, the insurance company has found that the probabilities for these events are 0.20 for minor damage, 0.10 for major damage, 0.07 for video damage, and 0.05 for replacement.

   (a) What proportion of policies receive no payments?

   (b) What premium should the insurance company charge to have an expected net gain of \$100 per policy?

   (c) What is the s.d. of the insurance company's net gain?

   Solution. This is similar to the previous problem. Let X denote the payout the insurance company makes to a policy holder, and let Y denote the net gain to the insurance company of a policy. Then

$$Y = \text{Premium} - X,$$

and so

$$E(Y) = \text{Premium} - E(X),$$

$$Var(Y) = Var(X).$$

The p.m.f. of X is given above,

$$P(X = 100) = 0.20, \;\; P(X = 500) = 0.10, \;\; P(X = 800) = 0.07, \;\; P(X = 1200) = 0.05.$$

Then $P(X = 0) = 0.58$, and

$$E(X) = (100)(0.20) + (500)(0.10) + (800)(0.07) + (1200)(0.05) = 186,$$

$$E(X^2) = (100^2)(0.20) + (500^2)(0.10) + (800^2)(0.07) + (1200^2)(0.05) = 143800.$$

Hence,

$$Var(Y) = 143800 - 186^2 = 109204,$$

and

$$sd(Y) = \sqrt{109204} = 330.46.$$

The premium required to have an expected net gain of \$100 would be

$$100 + E(X) = 286.$$

Homework 4

1. Remote clients on a distributed network generate calls to a file server randomly and independently at a mean rate of 6 calls per hour.

   (a) What is the probability that there will be no more than 3 calls generated during 1 hour?

   (b) What is the probability that exactly 6 calls are generated during 1 hour?

   (c) Suppose that the network will be expanded so that after this expansion calls to the file server will arrive at a mean rate of 24 calls per hour. Also suppose that the current file server will be overloaded if it receives more than 2 calls during 1 minute. What is the probability the server will be overloaded?

2. A communications center has one link that is open 40% of the time. Suppose the center polls this link every minute to determine whether or not the link is open, and suppose that whether or not it is open at one time is independent of whether or not it is open any other time.

   (a) What is the probability that the link will be busy (not open) each of the first 5 times it is polled?

   (b) What is the probability that the link will be open at least 4 times during the first 15 pollings?

   (c) What is the probability that it will take more than 3 pollings before finding an open time?

3. Suppose that 10% of switches sold by a vendor are defective. If a company purchases 15 switches from this vendor, what is the probability that

   (a) no more than 2 switches are bad;

   (b) between 3 and 5 switches, inclusively, are bad?

   (c) Suppose that these switches will be used for 12 subnets in this company's network, with the extra switches used for backups. What is the probability that all subnets will have non-defective switches and there will be at least 1 non-defective backup switch?

4. The server for a corporation's web site experiences a critical overload at times that occur randomly and independently at an average rate of 2 per month.

   (a) What is the probability that exactly 4 overloads occur during one month? more than 5 occur during one month.

   (b) What would you say if more than 8 critical overloads occur during a **two** month period of time?

5. Identical computer components are shipped in boxes of 5 from a manufacturing site that has a 10% defective rate. Boxes are tested in random order.

   (a) What is the probability that a randomly selected box has only non-defective components?

   (b) What is the probability that at least 8 of 10 randomly selected boxes have only non-defective components?

   (c) What is the expected value and standard deviation of the number of boxes tested until a box without defective components is found?

6. A particular genetic marker occurs in just 0.2% of people who are native south Asians. In a group of 2000 such people, assumed to be randomly selected, what is the probability that

   (a) no more than 2 will have that marker;

   (b) between 4 and 8, inclusively, will have the marker;

   (c) none will have the marker.

Solutions for Homework 4

1. Use Poisson model.

   (a) Mean for 1 hour is 6. $P(N \leq 3) = 0.1512$.

   (b) $P(N = 6) = 0.1606$.

   (c) Mean for 1 minute is $24/60 = 0.40$.

   $$P(\text{overload}) = P(N > 2) = 1 - P(N \leq 2) = 1 - 0.9921 = 0.0079.$$

2. These are independent Bernoulli trials with $P(\text{open}) = 0.40$.

   (a) The probability that the link will be busy (not open) each of the first 5 times it is polled is
   $.6^5 = 0.0778$.

   (b) Use Binomial(15,.4).

   $$P(N \geq 4) = 1 - P(N \leq 3) = 1 - 0.0905 = 0.9095.$$

   (c) This event is equivalent to the event that the link is busy for each of the first 3 pollings,

   $$0.6^3 = 0.216.$$

3. Use Binomial(15,.10) for number of defective switches among the 15 purchased.

(a) $P(N \leq 2) = 0.8160$.

(b) $P(3 \leq N \leq 5) = 0.1818$.

(c) Need to find probability there are at least 13 good switches among the 15 purchased. This corresponds to the event that there are no more than 2 defective switches, $P(N \leq 2) = 0.8160$.

4. Use Poisson model.

   (a) $P(N = 4) = 0.0902$.
   $P(N > 5) = 1 - P(N \leq 5) = 1 - 0.9834 = .0166$.

   (b) Mean for two months is 4.

   $$P(N > 8) = 1 - P(N \leq 8) = 1 - 0.9786 = 0.0214.$$

   It is highly unlikely (2% chance) that this would occur by chance, so this number is unually high.

5. The number of defectives in a box has a Binomial(n=5,p=0.10) distribution.

   (a) $P(N = 0) = .9^5 = 0.5905$.

   (b) The number of boxes with only non-defective components out of 10 randomly selected boxes is Binomial(n=10,p=0.5905). Hence,

   $$P(M \geq 8) = 0.1525.$$

   (c) Geometric(p=0.5905),

   $$P(X = k) = p(1 - p)^{k-1}, \ k \geq 1,$$

   where $p = 0.5905$.

6. The number of people with this marker has a binomial distribution with $n = 2000$ and $p = 0.002$, so n is large and p is small. Use Poisson to approximate this binomial with

   $$\lambda = np = (2000)(.002) = 4.$$

   Use Table II on p. 569.

   (a)

   $$P(N \leq 2) = .0183 + .0733 + .1465 = .2381.$$

   (b)

   $$P(4 \leq N \leq 8) = .1954 + .1563 + .1042 + .0595 + .0298 = .5452.$$

   (c)

   $$P(N = 0) = .0183.$$

Homework 5

The solutions for Exercise 6.20, Exercise 6.41, Exercise 6.53, Exercise 6.59, Exercise 6.70 with detailed explanations will be presented here at 5:30 pm CST on Tuesday, Nov. 22 instead of in class. Submit any questions you have about these solutions to me via email. Generally relevant questions and their answers will be posted here as well. The remainder of these problems are due on Nov. 29.

1. Exercise 6.20, p. 208. **Hint**: use the transformation $y = x^2$.
2. Exercise 6.23, p. 209. **Hint**: use the transformation $y = x^\beta$.
3. Exercise 6.41, p. 220
4. Exercise 6.42, p. 220
5. Exercise 6.52, p. 229
6. Exercise 6.53, p. 229
7. Exercise 6.59, p. 229
8. Exercise 6.60, p. 229
9. Exercise 6.70, p. 231. Additional question: what is the upper 5% of reduction in oxygen consumption during periods of TM?
10. Exercise 6.72, p. 231
11. Exercise 6.78, p. 231
12. Exercise 6.80, p. 232
13. Use moment generating functions for the following:

    (a) Let $X_1, \cdots, X_n$ be independent r.v.'s and let the distribution of X_k be normal with mean μ_k and variance σ^2, $1 \le k \le n$ (same variance, different means). Find the distribution of

    $$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

    (b) Let $X_1, \cdots, X_n$ be independent, identically distributed r.v.'s each of which has a gamma distribution with parameters α, β. Find the distribution of

    $$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

    Find the mean and s.d. of $\bar{X}$.

Solutions for Homework 5

Exercise 6.20

1.

$$\begin{aligned}\mu &= \int_0^\infty x(2\alpha)xe^{-\alpha x^2}dx \\ &= \int_0^\infty 2\alpha x^2 e^{-\alpha x^2}dx\end{aligned}$$

Let $y = x^2$. Then

$$x = y^{1/2} \text{ and } dx = 0.5y^{-1/2}.$$

This gives,

$$\mu = \alpha \int_0^\infty y^{1/2}e^{-\alpha y}dy.$$

Note that this integral is a generalized gamma integral with shape $3/2$ and scale $1/\alpha$. Therefore,

$$\begin{aligned}\mu &= \frac{\alpha\Gamma(3/2)}{\alpha^{3/2}} \\ &= \frac{1}{2}\Gamma(1/2)\alpha^{-1/2}.\end{aligned}$$

Since $\Gamma(1/2) = \sqrt{\pi}$, then

$$\mu = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}.$$

2. Use the same transformation to obtain the second moment.

$$\begin{aligned}E(X^2) &= \int_0^\infty 2\alpha x^3 e^{-\alpha x^2}dx \\ &= \alpha \int_0^\infty y^{3/2}e^{-\alpha y}y^{-1/2}dy \\ &= \alpha \int_0^\infty ye^{-\alpha y}dy \\ &= \alpha\Gamma(2)\alpha^{-2} \\ &= \alpha^{-1}.\end{aligned}$$

Therefore,

$$\sigma^2 = \alpha^{-1} - \frac{1}{4}\frac{\pi}{\alpha} = \frac{1}{\alpha}\left(1 - \frac{\pi}{4}\right).$$

Exercise 6.23. To find k, we must evaluate the integral,

$$\int_0^\infty x^{\beta-1}e^{-\alpha x^\beta}\,dx.$$

Note that the integrand has the form $e^U dU$ up to a constant, where $U = -\alpha x^\beta$. Therefore,

$$\begin{aligned}\int_0^\infty x^{\beta-1}e^{-\alpha x^\beta}\,dx &= -\frac{1}{\alpha\beta}e^{-\alpha x^\beta}\Big|_0^\infty. \\ &= \frac{1}{\alpha\beta}.\end{aligned}$$

This implies that $k = \alpha\beta$. To obtain μ, we can use the transformation,

$$y = x^\beta$$

Then

$$dx = \frac{1}{\beta}y^{1/\beta-1}dy$$

and so

$$\begin{aligned}\mu &= \alpha\beta\int_0^\infty x^\beta e^{-\alpha x^\beta}dx \\ &= \alpha\int_0^\infty ye^{-\alpha y}y^{1/\beta-1}dy \\ &= \alpha\int_0^\infty y^{1/\beta}e^{-\alpha y}dy\end{aligned}$$

Note that this is a generalized gamma integral with shape $1+1/\beta$ and scale $1/\alpha$. Hence,

$$\begin{aligned}\mu &= \frac{\alpha\Gamma(1+1/\beta)}{\alpha^{1+1/\beta}} \\ &= \alpha^{-1/\beta}\Gamma(1+1/\beta).\end{aligned}$$

Exercise 6.41. MGF of Poisson is

$$M(t) = Ee^{tX} = exp\{\lambda(e^t-1)\},$$

and so, MGF of Z is

$$\begin{aligned}M_Z(t) &= E\exp\{t\frac{X-\lambda}{\sqrt{\lambda}}\} \\ &= e^{-t\sqrt{\lambda}}M(t/\sqrt{\lambda}) \\ &= \exp\{-t\sqrt{\lambda}+\lambda(e^{t/\sqrt{\lambda}}-1)\}.\end{aligned}$$

The logarithm of this MGF is

$$\log(M_Z(t)) = -t\lambda^{1/2} + \lambda\left(e^{t/\sqrt{\lambda}} - 1\right).$$

Taylor series approximation of

$$(e^x - 1)$$

as $x \to 0$ is

$$(e^x - 1) = x + \frac{1}{2}x^2 + O(x^3),$$

where $O(x^3)$ satisfies

$$\sup_{x>0} \frac{O(x^3)}{x^3} < \infty.$$

Apply this approximation to $\log(M_Z(t))$ by replacing x with $t/\sqrt{\lambda}$:

$$\begin{aligned}
\log(M_Z(t)) &= -t\lambda^{1/2} + \lambda(e^{t/\sqrt{\lambda}} - 1) \\
&= -t\lambda^{1/2} + \lambda(t\lambda^{-1/2} + \frac{t^2}{2}\lambda^{-1} + O(\lambda^{-3/2})) \\
&= \frac{t^2}{2} + O(\lambda^{-1/2}).
\end{aligned}$$

Therefore,

$$\lim_{\lambda\to\infty} \log(M_Z(t)) = \frac{t^2}{2},$$

which is the logarithm of the MGF of the standard normal distribution.

Exercise 6.42. MGF of $Gamma(\alpha, \beta)$ is

$$Ee^{tX} = (1 - \beta t)^{-\alpha}$$

The mean and s.d. are $\alpha\beta$ and $\beta\sqrt{\alpha}$, so standardized gamma r.v. is

$$Z = \frac{X - \alpha\beta}{\beta\sqrt{\alpha}}$$

MGF of Z is

$$M_Z(t) = E\exp\{tZ\} = e^{-t\sqrt{\alpha}}E\exp\{\frac{t\sqrt{\alpha}}{\beta}X\} = e^{-t\sqrt{\alpha}}(1 - t/\sqrt{\alpha})^{-\alpha}.$$

Take logarithm of both sides.

$$\log M_Z(t) = -t\sqrt{\alpha} - \alpha\log(1 - t/\sqrt{\alpha}).$$

Taylor series approximation for $\log(1-x)$ as $x \to 0$ is

$$\log(1-x) = -x - \frac{1}{2}x^2 + O(x^3).$$

Apply this with $x = t/\sqrt{\alpha}$ to obtain

$$\log M_Z(t) = -t\sqrt{\alpha} + t\sqrt{\alpha} + \frac{1}{2}t^2 + O(\alpha^{-1/2}) \to \frac{1}{2}t^2$$

as $\alpha \to \infty$, which is the logarithm of the MGF of the standard normal distribution.

Exercise 6.52. Profit is gross sales minus sales cost,

$$\text{Profit} = X - 8n,$$

$$E(\text{Profit}) = E(X) - 8n = 160\sqrt{n} - 8n.$$

To find maximum, set derivative equal to 0:

$$160(1/2)n^{-1/2} - 8 = 0,$$

so $n = 10^2 = 100$.

Exercise 6.53. We must find $P(X > 12)$ when X has Gamma distribution with parameters $\alpha = 3$ and $\beta = 2$. This is the integral

$$P(X > 12) = \frac{1}{\Gamma(3)2^3}\int_{12}^{\infty} x^2 e^{-x/2}dx = \frac{1}{16}\int_{12}^{\infty} x^2 e^{-x/2}dx.$$

It is simpler to first substitute $y = x/2$ to give

$$P(X > 12) = \frac{1}{2}\int_{6}^{\infty} y^2 e^{-y}dy.$$

Apply integration by parts twice to this integral.

$$\begin{aligned} P(X > 12) &= \left[-\frac{1}{2}y^2 e^{-y}\right]_{y=6}^{\infty} + \frac{1}{2}\int_6^{\infty} ye^{-y}dy \\ &= \frac{1}{2}6^2 e^{-6} + \int_6^{\infty} ye^{-y}dy \\ &= 18e^{-6} + \left[-\frac{1}{2}ye^{-y}\right]_6^{\infty} + \int_6^{\infty} e^{-y}dy \\ &= 18e^{-6} + 6e^{-6} + e^{-6} \\ &= 25e^{-6} \end{aligned}$$

Exercise 6.59.

$$\begin{aligned} P(X < .10) &= \frac{\Gamma(11)}{\Gamma(2)\Gamma(9)}\int_0^{.1} x(1-x)^8 dx \\ &= \frac{10!}{1!8!}\int_0^{.1} x(1-x)^8 dx. \end{aligned}$$

We can either use integration by parts or substitute $y = 1 - x$ to evaluate this integral.

$$\begin{aligned} P(X < .10) &= 90\int_{.9}^{1}(1-y)y^8dy \\ &= 90\int_{.9}^{1}y^8dy - 90\int_{.9}^{1}y^9dy \\ &= 10 - .9^9 - 9 + .9^{10} \\ &= 1 - (.1)(.9^9). \end{aligned}$$

Exercise 6.60.

$$E(X) = \frac{\alpha}{\alpha+\beta} = \frac{1}{5}.$$

$$\begin{aligned} P(X \geq .25) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_{.25}^{1}x^{1-1}(1-x)^{4-1}dx \\ &= -\frac{\Gamma(5)}{4\Gamma(1)\Gamma(4)}(1-x)^4\Big|_{.25}^{1} \\ &= (1-.25)^4 \\ &= 81/256 \\ &= 0.3164. \end{aligned}$$

Exercise 6.70. Find areas under the normal density. First convert to z-scores by subtracting the mean and then dividing by the s.d. Then use Table III on p. 574 of the text.

1.

$$P(X \geq 44.5) = P(Z \geq \frac{44.5 - 37.6}{4.6}) = P(Z \geq 1.5).$$

Table gives area between 0 and z which in this case for $z = 1.5$ is 0.4332. We need the area to the right of 1.5, which is

$$0.50 - 0.4332 = .0668.$$

2.

$$P(X \leq 35) = P(Z \leq \frac{35 - 37.6}{4.6} = P(Z \leq -0.57).$$

Because of the symmetry of the normal curve, this area is the same as

$$P(X \leq 35) = P(Z \leq -0.57) = P(Z \geq 0.57) = 0.5 - 0.2157 = .2843.$$

3.

$$P(30 < X < 40) = P(-1.65 < Z < 0.52) = 0.4505 + 0.1985 = 0.6490.$$

4. Upper 5% is defined by a value such that only 5% of the time reduction exceeds thar value. We first answer this question for the standard normal distribution and then convert that answer to the normal distribution for reduction in oxygen consumption. So we must find the z-value such that the area to the right of z is 0.05. Then the area between 0 and z must be 0.45 and this is the **area** we must search for in the table. The closest we can come to 0.45 is 0.4495 ($z = 1.64$) or 0.4505 ($z = 1.65$). Either value can be used. For $z = 1.65$ we have

$$x = 37.6 + 1.65 * 4.6 = 45.19$$

So the upper 5% of reduction in oxygen consumption is 45.19.

Exercise 6.72. We need to satisfy

$$0.03 = P(X < 6) = P(Z < (6 - \mu)/0.05).$$

Since $P(Z < -1.88) = 0.0301$, then

$$-1.88 = (6 - \mu)/0.05$$

which has solution $\mu = 6.094$.

Exercise 6.78. Number with bad side effects has a binormial distribution with $n = 120$ and $p = 0.23$. This can be approximated by normal distribution with mean $np = 27.6$ and s.d.

$$s.d. = \sqrt{npq} = \sqrt{120(0.23)(0.77)} = 4.61$$

This gives

$$\begin{aligned} P(X > 32) &\approx P(Z \geq \frac{32.5 - 27.6}{4.61} \\ &= P(Z \geq 1.06) \\ &= 1 - 0.8554 \\ &= 0.1446. \end{aligned}$$

Exercise 6.80. Note that proportion of heads has approximately a normal distribution with mean $p = 0.50$ and s.d. $\sqrt{pq/n} = 0.50/\sqrt{n}$.

1. For $n = 100$,

$$\begin{aligned} P(0.49 \leq X \leq 0.51) &\approx P((0.49 - 0.50)/0.05 \leq Z \leq (0.51 - 0.50)/0.05) \\ &= P(-0.2 \leq Z \leq 0.20) \\ &= 0.5793 - 0.4207 \\ &= 0.1586 \end{aligned}$$

2. For $n = 1000$,

$$\begin{aligned} P(0.49 \le X \le 0.51) &\approx P((0.49 - 0.50)/0.0158 \le Z \le (0.51 - 0.50)/0.0158) \\ &= P(-0.63 \le Z \le 0.63 \\ &= 0.7357 - 0.2643 \\ &= 0.4714 \end{aligned}$$

3. For $n = 10000$,

$$\begin{aligned} P(0.49 \le X \le 0.51) &\approx P((0.49 - 0.50)/0.005 \le Z \le (0.51 - 0.50)/0.005) \\ &= P(-2.00 \le Z \le 2.00 \\ &= 0.9772 - 0.0228 \\ &= 0.9544 \end{aligned}$$

Let $X_1, \cdots, X_n$ be independent r.v.'s and let the distribution of X_k be normal with mean μ_k and variance σ^2, $1 \le k \le n$ (same variance, different means). Find the distribution of

$$\bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k.$$

Solution:

$$\begin{aligned} Ee^{t\bar{X}} &= E \exp\{\frac{t}{n} \sum_{k=1}^{n} X_k\} \\ &= \prod_{k=1}^{n} E \exp\{\frac{t}{n} X_k\} \\ &= \prod_{k=1}^{n} \exp\{\frac{t}{n}\mu_k + \frac{t^2\sigma^2}{2n^2}\} \\ &= \exp\{t\bar{\mu} + \frac{t^2}{2}\frac{\sigma^2}{n}\}, \end{aligned}$$

where

$$\bar{\mu} = \frac{1}{n} \sum_{k=1}^{n} \mu_k.$$

This is the MGF of the normal distribution with mean $\bar{\mu}$ and variance σ^2/n.

Let $X_1, \cdots, X_n$ be independent, identically distributed r.v.'s each of which has a gamma distribution with parameters α, β. Find the distribution of

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Find the mean and s.d. of $\bar{X}$.
Solution:

$$\begin{aligned} Ee^{t\bar{X}} &= E\exp\{\sum_{k=1}^{n}\frac{t}{n}X_k\} \\ &= \prod_{k=1}^{n}(1-\beta t/n)^{-\alpha} \\ &= \left(1-\frac{\beta}{n}t\right)^{-n\alpha}. \end{aligned}$$

This is the MGF of the Gamma distribution with shape $n\alpha$ and scale β/n. Hence,

$$E\bar{X} = n\alpha\beta/n = \alpha\beta$$

and

$$SD(\bar{X}) = \sqrt{n\alpha}\beta/n = \frac{\sqrt{\alpha}\beta}{\sqrt{n}}.$$