MH 2500 Probability and introduction to statistics

Tutorial Week 2. Chapter 1: Probability. Problems

1. **(1.13c)** In a game of poker, what is the probability that a five-card hand will contain a full house (three cards of one value and two cards of another value)?

2. **(1.30)** A group of 60 second graders is to be randomly assigned to two classes of 30 each. Five of the second graders, Marcelle, Sarah, Michelle, Katy, and Camerin, are close friends.

   (a) What is the probability that they will all be in the same class?

   (b) What is the probability that exactly four of them will be in the same class?

   (c) What is the probability that Marcelle will be in one class and her friends in the other?

3. **(1.34)** Consider the following important combinatorial identity.

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

   (a) Prove the identity by interpreting its meaning combinatorially. *Hint: Consider an urn with m balls of which n are white and $m-n$ are black.*

   (b) Prove the identity algebraically. *Hint: Consider* $(1+a)^n (1+a)^{m-n} = (1+a)^m$. *Now use the Binomial Theorem and match coefficients.*

4. **(1.39)** A monkey at a typewriter types each of the 26 letters of the alphabet exactly once, the order being random.

   (a) What is the probability that the word "Hamlet" appears somewhere in the string of letters?

MH 2500 Probability and introduction to statistics

Tutorial Week 2. Chapter 1: Probability. Solutions

1. **(1.13c)** In a game of poker, what is the probability that a five-card hand will contain a full house (three cards of one value and two cards of another value)?
   Let A be the event that a (randomly chosen) poker hand is a "Full House". We are interested in finding $\Pr(A)$. Since the cards are selected randomly, all outcomes/poker hands are equally likely, and so it is enough to count outcomes/poker hands, i.

   $$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{\text{Number of poker hands with "Full House"}}{\text{Number of poker hands}}$$

   Since the draw is done without replacement and without respect to order it is the technique of combinations that should be used to count outcomes. Without any restrictions it is clear that

   $$|\Omega| = C_5^{52} = \binom{52}{5} = 2\,598\,960$$

   The number of outcomes in the numerator is a little bit more tricky. This can be done in different ways. First we choose the two values that are to constitute the full house, which can be done in $\binom{13}{2}$ ways. Then we need to pick which of the two chosen values that should be the three of a kind, which can be done in $\binom{2}{1}$ ways. We note that these two steps also can be done more logically using permutations since $\binom{13}{2}\binom{2}{1} = P_2^{13}$. There are four cards of each value which means that the three of a kind can be chosen in $\binom{4}{3}$ ways, and the pair can be chosen in $\binom{4}{2}$ ways. Finally, for completeness, we are from the remaining 44 cards not to choose any cards which can be done in $\binom{44}{0} = 1$ way. It follows that

   $$|A| = \binom{13}{2}\binom{2}{1}\binom{4}{3}\binom{4}{2}\binom{44}{0} = 3\,744$$

   and so we conclude that

   $$\Pr(A) = \frac{\binom{13}{2}\binom{2}{1}\binom{4}{3}\binom{4}{2}\binom{44}{0}}{\binom{52}{5}} = \frac{3\,744}{2\,598\,960} = \frac{6}{4\,165} \approx 0.00144$$

2. **(1.30)** First we let

$$A_k = k \text{ of the friends end up in the first class}$$

for $k = 0, 1, 2, 3, 4, 5$. Since the draw is done without replacement and without respect to order, it follows that we are to use the concept of combinations for the probability assessment, i.e. it follows that

$$\Pr(A_k) = \frac{\binom{5}{k}\binom{55}{30-k}}{\binom{60}{30}}, \quad k = 0, 1, 2, 3, 4, 5$$

At this point it is important to realize that the A_k:s are all pairwise disjoint.

(a) We now have that

$$\begin{aligned}\Pr(\text{All five friends in the same class}) &= \Pr(A_0 \cup A_5) = \Pr(A_0) + \Pr(A_5) \\ &= \frac{\binom{5}{0}\binom{55}{30-0} + \binom{5}{5}\binom{55}{30-5}}{\binom{60}{30}} = 0.052\end{aligned}$$

(b) Reasoning in a similar way we find that

$$\Pr(\text{Exactly four of the friends in the same class}) =$$
$$= \Pr(A_1 \cup A_4) = \Pr(A_1) + \Pr(A_4) = \frac{\binom{5}{1}\binom{55}{30-1} + \binom{5}{4}\binom{55}{30-4}}{\binom{60}{30}} = 0.301$$

(c) In the previous problem we found the probability that exactly four of the friends end up in the same class, i.e. that exactly one of the friends will be alone (so to say). Now we make some observations. If we let B_1 be the event that (only) Marcelle is alone, B_2 the event that (only) Sarah is alone, and so on, it is clear that $B_1, B_2, \ldots, B_5$ are pairwise disjoint and, by symmetry, equiprobable events. Furthermore, it is clear that

$$B_1 \cup B_2 \cup \cdots \cup B_5 = A_1 \cup A_4$$

It therefore follows that

$$\Pr(B_1) = \frac{1}{5}\Pr(A_1 \cup A_4) = \frac{1}{5} \cdot \frac{\binom{5}{1}\binom{55}{30-1} + \binom{5}{4}\binom{55}{30-4}}{\binom{60}{30}} = 0.06$$

Alternatively, we can see this problem as drawing from an urn with *three* types of objects; Marcelle, the four other friends, and the others. This leads to the calculation

$$\Pr(\text{Marcelle alone}) = \frac{\binom{1}{1}\binom{4}{0}\binom{55}{29} + \binom{1}{0}\binom{4}{4}\binom{55}{26}}{\binom{60}{30}} = 0.06$$

with the same result.

3. **(1.34)** Consider the following important combinatorial identity.

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

(a) Prove the identity by interpreting its meaning combinatorially. *Hint: Consider an urn with m balls of which n are white and $m-n$ are black.*
In what follows, we assume that the draw is done without replacement and without respect to order. The right hand side of the expression represents the number of ways to draw n balls from an urn containing m balls. In the left hand side of the expression, the term $\binom{n}{k}\binom{m-n}{n-k}$ represents the number of ways to draw n balls from an urn containing m balls, such that exactly k of the drawn balls are white (and apparantly $n-k$ balls are black). When summing such terms for every possible value of k, we should of course get the same answer as if we did not having this restriction.

(b) Prove the identity algebraically. *Hint: Consider* $(1+a)^n (1+a)^{m-n} = (1+a)^m$. *Now use the Binomial Theorem and match coefficients.*
The hint gives us that

$$(1+a)^n (1+a)^{m-n} = \sum_{i=0}^{n} \binom{n}{i} a^i \sum_{j=0}^{m-n} \binom{m-n}{j} a^j = \sum_{i=0}^{n} \sum_{j=0}^{m-n} \binom{n}{i} \binom{m-n}{j} a^{i+j}$$

which means that the coefficient of a^n is

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

For the right hand side a similar approach yields

$$(1+a)^m = \sum_{k=0}^{m} \binom{m}{k} a^k$$

which means that the coefficient of a^n is $\binom{m}{n}$. Since the coefficient of a^n must be the same for the left hand side and the right hand side the statement follows.

4. **(1.39)** A monkey at a typewriter types each of the 26 letters of the alphabet exactly once, the order being random.

   (a) What is the probability that the word "Hamlet" appears somewhere in the string of letters?
   Let A be the event that a (randomly chosen) string of letters contains the word "Hamlet". We are interested in finding $\Pr(A)$. Since the order of the letters are random, all outcomes/letter strings are equally likely, i.e. it is enough to count outcomes/letter strings, i.e.

$$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{\text{Number of letter strings containing the word "Hamlet"}}{\text{Number of letter strings}}$$

   Since the draw is done without replacement and with respect to order it is the technique of permutations that should be used to count outcomes. Without any restrictions it is clear that

$$|\Omega| = P_{26}^{26} = 26! \approx 4.0329 \times 10^{26}$$

   In order to find $|A|$ we first pick the letters HAMLET and put them in this order. The remaining 20 letters can be permuted in $P_{20}^{20} = 20!$ ways. Furthermore, there are 21 starting positions for the word HAMLET which means that

$$|A| = 21 \cdot P_{20}^{20} = 21 \cdot 20! \approx 5.1091 \times 10^{19}$$

   and so we conclude that

$$\Pr(A) = \frac{21 \cdot 20!}{26!} = \frac{21!}{26!} = \frac{1}{7\,893\,600}$$

MH 2500 Probability and introduction to statistics

Tutorial 2 Week 3. Chapter 1: Probability. Problems

1. (**1.60**) If B is an event, with $\Pr(B) > 0$, show that the set function $Q(A) = \Pr(A \mid B)$ satisfies the axioms for a probability measure. Thus, for example,

$$\Pr(A \cup C \mid B) = \Pr(A \mid B) + \Pr(C \mid B) - \Pr(A \cap C \mid B)$$

2. On each turn in the dice game Yahtzee, a player gets up to three rolls of the dice. He or she can save any dice that are wanted to complete a combination and then re-roll the other dice. After the third roll, the player must find a place to put the score (though he or she can choose to end the turn and score after one or two rolls, if desired). Suppose you are at the end of the game, and the only combination that is left is "Yahtzee", which means that all five dice must show the same number of pips for you to get points. In your first roll you succeed in getting three fours which you save. Determine the probability that you will manage to get "Yahtzee".
Hint: Use the Law of Total Probability.
Remark: If you are not familiar with the game Yahtzee, I suggest that you take a look at http://en.wikipedia.org/wiki/Yahtzee.

3. (**1.76**) Suppose that n components are connected in series. For each unit, there is a backup unit, and the system fails if and only if both a unit and its backup fail. Assuming that all the units are independent and fail with probability p, what is the probability that the system works? For $n = 10$ and $p = 0.05$, compare these results with those of Example F in Section 1.6.

4. Person A and Person B take turns rolling a pair of dice, and A starts. The game ends if either the sum of pips rolled by A is nine, or if the sum of pips rolled by B is six. Determine the probability that the last roll of the pair of dice is made by A.
Hint: Use ideas from the solution to the Craps problem discussed in the Monday lecture.

MH 2500 Probability and introduction to statistics

Tutorial 2 Week 3. Chapter 1: Probability. Solutions

1. (**1.60**) If B is an event, with $\Pr(B) > 0$, show that the set function $Q(A) = \Pr(A \mid B)$ satisfies the axioms for a probability measure. Thus, for example,

$$\Pr(A \cup C \mid B) = \Pr(A \mid B) + \Pr(C \mid B) - \Pr(A \cap C \mid B)$$

We show that each of the three Kolmogorov axioms are satisfied.

- $\Pr(\Omega) = 1$.

$$Q(\Omega) = \Pr(\Omega \mid B) = \frac{\Pr(\Omega \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1$$

- Let $A \subset \Omega$. Then it hold that $\Pr(A) \geq 0$, and so

$$Q(A) = \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Since both $A \subset \Omega$ and $B \subset \Omega$ we have that $A \cap B \subset \Omega$ and therefore $\Pr(A \cap B) \geq 0$. Furthermore, $\Pr(B) > 0$ and so $Q(A) \geq 0$.

- Let $A_1 \subset \Omega$ and $A_2 \subset \Omega$ be events such that $A_1 \cap A_2 = \emptyset$. Then it follows that

$$\begin{aligned} Q(A_1 \cup A_2) &= \Pr(A_1 \cup A_2 \mid B) = \frac{\Pr((A_1 \cup A_2) \cap B)}{\Pr(B)} = \\ &= \frac{\Pr((A_1 \cap B) \cup (A_2 \cap B))}{\Pr(B)} = \frac{\Pr(A_1 \cap B) + \Pr(A_2 \cap B)}{\Pr(B)} = \\ &= \Pr(A_1 \mid B) + \Pr(A_2 \mid B) = Q(A_1) + Q(A_2) \end{aligned}$$

since $A_1 \cap B$ and $A_2 \cap B$ are disjoint.

2. On each turn in the dice game Yahtzee, a player gets up to three rolls of the dice. He or she can save any dice that are wanted to complete a combination and then re-roll the other dice. After the third roll, the player must find a place to put the score (though he or she can choose to end the turn and score after one or two rolls, if desired). Suppose you are at the end of the game, and the only combination that is left is "Yahtzee", which means that all five dice must show the same number of pips for you to get points. In your first roll you succeed in getting three fours which you save. Determine the probability that you will manage to get "Yahtzee". *Hint: Use the Law of Total Probability.*
Remark. There are more than one way to solve this problem. We are interested in finding the probability for the event

$$A = \text{Yahtzee (given three fours in the first roll)}$$

In order to find $\Pr(A)$ we make use of the hint by applying the law of total probability. The first thing that we need to do is to find an appropriate partition of the sample space. Since we have already saved the three fours from the first roll, we will only cast two dice in the second roll and therefore an appropriate partition seems to be based on the number of fours in the second roll, i.e.

$$\begin{aligned} B_0 &= \text{No four in the second roll} \\ B_1 &= \text{Exactly one four in the second roll} \\ B_2 &= \text{Two fours in the second roll} \end{aligned}$$

where we should note the "natural" notation that has been used. It is not hard to confirm that this really is a partition of the sample space. If we now let the events D_1 and D_2 represent that we get fours on dice 1 and dice 2, respectively, it is clear that $\Pr(D_1) = \Pr(D_2) = 1/6$ and that D_1 and D_2 are independent. Therefore

$$\begin{aligned} \Pr(B_0) &= \Pr(D_1^c \cap D_2^c) = \Pr(D_1^c)\Pr(D_2^c) = \frac{5}{6} \cdot \frac{5}{6} = \frac{25}{36} \\ \Pr(B_1) &= \Pr[(D_1 \cap D_2^c) \cup (D_1^c \cap D_2)] = \Pr(D_1 \cap D_2^c) + \Pr(D_1^c \cap D_2) = \\ &= \Pr(D_1)\Pr(D_2^c) + \Pr(D_1^c)\Pr(D_2) = \frac{1}{6} \cdot \frac{5}{6} + \frac{5}{6} \cdot \frac{1}{6} = \frac{10}{36} \\ \Pr(B_2) &= \Pr(D_1 \cap D_2) = \Pr(D_1)\Pr(D_2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} \end{aligned}$$

In the calculation of $\Pr(B_1)$ we used the fact that $D_1 \cap D_2^c$ and $D_1^c \cap D_2$ are disjoint. An alternative way to calculate $\Pr(B_1)$ is to use the fact that B_0, B_1, and B_2 is a partition of the sample space and therefore

$$\Pr(B_1) = 1 - [\Pr(B_0) + \Pr(B_2)] = 1 - \left(\frac{25}{36} + \frac{1}{36}\right) = \frac{10}{36}$$

The conditional probabilities for Yahtzee given these events are

$$\begin{aligned}
\Pr(A \mid B_0) &= \Pr(B_2) = \frac{1}{36} \\
\Pr(A \mid B_1) &= \Pr(\text{A four with } \textit{one} \text{ die}) = \frac{1}{6} \\
\Pr(A \mid B_2) &= 1
\end{aligned}$$

By the law of total probability it thus follows that

$$\begin{aligned}
\Pr(A) &= \Pr(B_0)\Pr(A \mid B_0) + \Pr(B_1)\Pr(A \mid B_1) + \Pr(B_2)\Pr(A \mid B_2) = \\
&= \frac{25}{36} \cdot \frac{1}{36} + \frac{10}{36} \cdot \frac{1}{6} + \frac{1}{36} \cdot 1 = \frac{121}{1296} \approx 0.0934
\end{aligned}$$

3. **(1.76)** Suppose that n components are connected in series. For each unit, there is a backup unit, and the system fails if and only if both a unit and its backup fail. Assuming that all the units are independent and fail with probability p, what is the probability that the system works? For $n = 10$ and $p = 0.05$, compare these results with those of Example F in Section 1.6. First, let

$$\begin{aligned}
A_k &= \text{The component in position } k \text{ works}, \quad k = 1, 2, \ldots, n \\
B_k &= \text{The backup unit in position } k \text{ works}, \quad k = 1, 2, \ldots, n
\end{aligned}$$

Now, let C be the event that the system works. Due to the fact that each components works (and fails) independently of each other, it follows that

$$\begin{aligned}
\Pr(C) &= \Pr((A_1 \cup B_1) \cap (A_2 \cup B_2) \cap \cdots \cap (A_n \cup B_n)) = \\
&= (\Pr(A_1 \cup B_1))^n = (\Pr(A_1) + \Pr(B_1) - \Pr(A_1)\Pr(B_1))^n = \\
&= \left(2(1-p) - (1-p)^2\right)^n = ((1-p)(1+p))^n
\end{aligned}$$

For the specific situation $n = 10$ and $p = 0.05$, this probability is

$$\Pr(C) = ((1-p)(1+p))^n = (0.95 \cdot 1.05)^{10} = 0.9753$$

4. Person A and Person B take turns rolling a pair of dice, and A starts. The game ends if either the sum of pips rolled by A is nine, or if the sum of pips rolled by B is six. Determine the probability that the last roll of the pair of dice is made by A. *Hint: Use ideas from the solution to the Craps problem discussed in the Monday lecture.*

   In each roll made by A, there are 4 outcomes (36,45, 54, and 63) for which the sum of pips is 9, and in each roll made by B, there are 5 outcomes (15, 24, 33, 42, and 51) for which the sum of pips is 6. Since there is a total of 36 outcomes in Ω and since all outcomes are equally likely to occur, the probability that A gets the sum nine in any given roll is $4/36 = 1/9$. Similarly, the probability that B gets the sum six in any given roll is $5/36$. Now consider the event

$$A_n = \text{Person } A \text{ wins in the } n\text{:th roll}$$

   In order for A to win in the n:th roll, we must have a situation where neither A nor B have already won (i.e. succeeded in getting nine and six, respectvely) in their first $n-1$ rolls. Since the results of different rolls are *independent*, it follows that

$$\Pr(A_n) = \left(1 - \frac{1}{9}\right)^{n-1} \left(1 - \frac{5}{36}\right)^{n-1} \cdot \frac{1}{9} = \left(\frac{62}{81}\right)^{n-1} \cdot \frac{1}{9}, \quad n = 1, 2, \ldots$$

   A wins if any of the events A_n occurs. Since these events are *pairwise disjoint*, we have that

$$\Pr(A) = \Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \Pr(A_n) = \frac{1}{9} \sum_{n=1}^{\infty} \left(\frac{62}{81}\right)^{n-1} = \frac{1}{9} \cdot \frac{1}{1 - \frac{62}{81}} = \mathbf{\frac{9}{19}}$$

MH 2500 Probability and introduction to statistics

Tutorial 3 Week 4. Chapter 2.1: Discrete Random Variables. Problems

1. (**1.10**) If n balls are distributed randomly into k urns, what is the probability that the last urn contains j balls? Solve this problem by first constructing a random variable.

2. (**1.30 revisited**) In the first tutorial session we discussed this problem. Take a look at the solutions presented on the course homepage.

   (a) In the solutions a collection of events is defined. Why, from a random variable point of view, is this collection logical?

   (b) Find the probability distribution of the random variable that gives rise to this collection of events.

   (c) In the last part of the problem we are to find the probability that Marcelle will be in one class and her friends in another. Explain why this problem cannot be solved using a random variable of any of the "common types".

3. Let $X \sim Ge(p)$.

   (a) (**2.19**) Find an expression for the cumulative distribution function of X.

   (b) (**2.21**) Show that

   $$\Pr(X > n - 1 + k \mid X > n - 1) = \Pr(X > k)$$

   In light of the construction of a geometric distribution from a sequence of independent bernoulli trials, how can this be interpreted so that it is "obvious"?

4. (**2.31**) Phone calls are received at a certain residence as a Poisson process with parameter $\lambda = 2$ per hour.

   (a) If Diane takes a 10 min shower, what is the probability that the phone rings during that time?

   (b) How long can her shower be if she wishes the probability of receiving no phone calls to be at least 0.5? *Remark. In the textbook it wrongly says "at most 0.5".*

MH 2500 Probability and introduction to statistics

Tutorial 3 Week 4. Chapter 2.1: Discrete Random Variables. Solutions

1. (**1.10**) If n balls are distributed randomly into k urns, what is the probability that the last urn contains j balls? Solve this problem by first constructing a random variable. If a drawn ball ends up in the last urn, we call it a "Success" and otherwise a "Failure". The sample size, i.e. the number of balls to be distributed, is predetermined. Since the balls are distributed randomly, the probability of "Success" is the same during the entire procedure. Since the number of urns is k, this constant probability of "Success" is $p = 1/k$. Finally, if we define our random variable as

$$X = \text{The number of balls that end up in the last urn}$$

it is clear that $X \sim Bin\,(n, 1/k)$. Therefore

$$p\,(j) = \binom{n}{j}\left(\frac{1}{k}\right)^{j}\left(1-\frac{1}{k}\right)^{n-j}, \quad j = 0, 1, 2, \ldots, n$$

2. (**1.30 revisited**) In the first tutorial session we discussed this problem. Take a look at the solutions presented on the course homepage.

   (a) In the solutions a collection of events is defined. Why, from a random variable point of view, is this collection logical? The collection

   $$A_k = k \text{ of the friends end up in the first class}$$

   for $k = 0, 1, 2, 3, 4, 5$, is clearly a partition of the sample space. Also, it is the partition induced by the random variable

   $$X = \text{The number of friends that end up in the first class}$$

   which is the logical choice of random variable for this situation.

   (b) Find the probability distribution of the random variable that gives rise to this collection of events. We are to draw 30 balls from an urn containing 60 balls. In this urn we have 5 "Successes" and 55 "Failures". The sample size, i.e. the number of balls to be distributed, is predetermined. Since we draw without replacement, the probability of "Success" will vary during the procedure. Finally, since we define our random variable as

   $$X = \text{The number of friends that end up in the first class}$$

it is clear that $X \sim Hyp\,(n = 30, r = 5, N = 60)$. Therefore

$$\Pr(X = k) = \Pr(A_k) = \frac{\binom{5}{k}\binom{55}{30-k}}{\binom{60}{30}}, \quad k = 0, 1, 2, 3, 4, 5$$

(c) In the last part of the problem we are to find the probability that Marcelle will be in one class and her friends in another. Explain why this problem cannot be solved using a random variable of any of the "common types". I we view this situation as an urn model, it is clear that the urn contains *three* types of objects; Marcelle, the four other friends, and the others. Since it is not an urn with only "Successes" and "Failures", none of the "common types" of random variables can be used.

3. Let $X \sim Ge\,(p)$.

(a) (**2.19**) Find an expression for the cumulative distribution function of X. Since $X \sim Ge\,(p)$, the pmf of X is given by

$$p(k) = (1-p)^{k-1}\,p, \quad k = 1, 2, \ldots$$

and so it follows that the cdf of X is given by

$$F(k) = \sum_{i=1}^{k} (1-p)^{i-1}\,p = p\sum_{i=0}^{k-1} (1-p)^{i} = \frac{p\left(1-(1-p)^k\right)}{1-(1-p)} = 1-(1-p)^k$$

(b) (**2.21**) Show that

$$\Pr(X > n-1+k \mid X > n-1) = \Pr(X > k)$$

In light of the construction of a geometric distribution from a sequence of independent bernoulli trials, how can this be interpreted so that it is "obvious"? In the a-part we found that

$$F(k) = 1-(1-p)^k, \quad k = 1, 2, \ldots$$

which means that

$$\begin{aligned}
\Pr(X > n-1+k \mid X > n-1) &= \frac{\Pr(X > n-1+k)}{\Pr(X > n-1)} = \frac{1-F(n-1+k)}{1-F(n-1)} = \\
&= \frac{1-\left(1-(1-p)^{n-1+k}\right)}{1-\left(1-(1-p)^{n-1}\right)} = \frac{(1-p)^{n-1+k}}{(1-p)^{n-1}} = \\
&= (1-p)^k = 1-F(k) = \Pr(X > k)
\end{aligned}$$

This is the co called “no memory” property of the geometric distribution. The interpretation of the conditional probability

$$\Pr(X > n - 1 + k \mid X > n - 1)$$

is that there is no “Success” among the first $n-1$ drawn balls, and that we want to determine the probability that there will be no “Success” among the next k drawn balls. Since every drawn ball is replaced, the knowledge that there is no “Success” among the first $n-1$ drawn balls is, from a probability point of view, the same as to say that the n:th ball do be drawn is the first ball of a seqence with the same properties as the original sequence.

4. (**2.31**) Phone calls are received at a certain residence as a Poisson process with parameter $\lambda = 2$ per hour.

   (a) If Diane takes a 10 min shower, what is the probability that the phone rings during that time? If we consider the random variable

   $$X = \text{The number of phone calls during a ten minute period}$$

   it follows that $X \sim Po(1/3)$. The probability that there will be at least one phone call during this time period is

   $$\Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - e^{-\frac{1}{3}} = 0.283$$

   (b) How long can her shower be if she wishes the probability of receiving no phone calls to be at least 0.5? *Remark. In the textbook it wrongly says “at most* 0.5*”.* Let t be the length of the shower (in fractions of an hour). If we let

   $$Y = \text{The number of phone calls during the shower}$$

   then $Y \sim Po(2t)$. Since the probability for no calls is given by

   $$\Pr(Y = 0) = e^{-2t}$$

   it is clear that we have to solve the equation

   $$e^{-2t} \geq \frac{1}{2} \iff -2t \geq -\ln 2 \iff t \leq \frac{\ln 2}{2} \approx 0.346\,57$$

   Hence, the shower should not be longer than $0.346\,57 \cdot 60 = 20.79$ minutes.

MH 2500 Probability and introduction to statistics

Tutorial 4 Week 5. Chapter 2.2: Continuous Random Variables. Problems

1. The pdf of the random variable X is given by

$$f(x) = \begin{cases} x, & 0 \le x \le 1 \\ c/x^3, & x > 1 \end{cases}$$

   (a) Determine the constant c.

   (b) Compute $\Pr(1/2 \le X \le 2)$.

2. (**2.42**) Find the pdf for the distance from an event to its nearest neighbor for a Poisson process in the plane.
   Remark. For such a process, the intensity parameter λ concerns the number of occurrences in an *area* of unit size.
   Hint. Let X represent the number of occurrences (i.e. neighbors) in a circle with radius r. Now let Y represent the distance to the nearest neighbor. What is the probability distribution of X? What about the relationship between X and Y?

MH 2500 Probability and introduction to statistics

Tutorial 4 Week 5. Chapter 2.2: Continuous Random Variables. Solutions

1. The pdf of the random variable X is given by

$$f(x) = \begin{cases} x, & 0 \le x \le 1 \\ c/x^3, & x > 1 \end{cases}$$

   (a) Determine the constant c. In order for f to be a pdf, it must hold that

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 x dx + c \int_1^{\infty} x^{-3} dx = \left[\frac{x^2}{2}\right]_0^1 + c\left[\frac{x^{-2}}{2}\right]_{\infty}^1 = \frac{1}{2} + \frac{c}{2}$$

   which means that $c = 1$.

   (b) Compute $\Pr(1/2 \le X \le 2)$.

$$\begin{aligned} \Pr(1/2 \le X \le 2) &= \int_{1/2}^1 x dx + \int_1^2 x^{-3} dx = \left[\frac{x^2}{2}\right]_{1/2}^1 + \left[\frac{x^{-2}}{2}\right]_2^1 = \\ &= \frac{1}{2} - \frac{1}{8} + \frac{1}{2} - \frac{1}{8} = \mathbf{\frac{3}{4}} \end{aligned}$$

2. Find the pdf for the distance from an event to its nearest neighbor for a Poisson process in the plane.
   Remark. For such a process, the intensity parameter λ concerns the number of occurrences in an *area* of unit size.
   Hint. Let X represent the number of occurrences (i.e. neighbors) in a circle with radius r. Now let Y represent the distance to the nearest neighbor. What is the probability distribution of X? What about the relationship between X and Y?
   We now focus on a certain occurence A. So, if we let

$$X = \text{The number of neighbors (to } A\text{) that are closer than } r$$

   it follows that $X \sim Po(\lambda\pi r^2)$. We are now interested in

$$Y = \text{The distance between } A \text{ and its nearest neighbor}$$

   If the distance to the nearest neighbor is greater than r, then there are no neighbors in an area of size $x = \pi r^2$. So

$$1 - F_Y(r) = \Pr(Y > r) = \Pr(X = 0) = e^{-\lambda\pi r^2}$$

which means that

$$F_Y(r) = 1 - e^{-\lambda\pi r^2}$$

and so the pdf of Y is given by

$$f_Y(r) = 2\lambda\pi r e^{-\lambda\pi r^2},\ r > 0$$

which means that Y follows a *Rayleigh distribution* (see PowerPoint 2.3, slide 3).

MH 2500 Probability and introduction to statistics

Tutorial 5 Week 6. Chapters 2.2-2.3, 4.1. Problems

1. The length of a human pregnancy can be considered to be a normally distributed random variable with a mean of 266 days and a standard deviation of 16 days.
   Remark 1. We solve this problem by using Table 2 of Appendix B.
   Remark 2. The random variable is *continuous*, and the values given should be interpreted as *exact* values, e.g. 240 days means $x = 240$.

   (a) Determine the probability that the length of the pregnancy for a randomly chosen pregnant woman will be less than 240 days, i.e. that it, roughly speaking, will last less than 8 months.

   (b) Determine the third quartile, q_3. Interpret this value.

   (c) Suppose that we study 5 randomly chosen pregnant women. Determine the probability that at least two of them will be "late", i.e. have a pregnancy that lasts longer than expected.

2. (**2.62**) A particle of mass m has a random velocity, V, which is normally distributed with parameters $\mu = 0$ and σ. Find the pdf of the kinetic energy, $E = \frac{1}{2}mV^2$ and try to identify the distribution of E as one of the commont types. *Hint.* $\sqrt{\pi} = \Gamma(1/2)$.

3. (**4.20**) Let $X \sim Po(\lambda)$. Find
$$E\left(\frac{1}{X+1}\right)$$

4. (**4.23**) A random square has a side length that is an exponential random variable. Find the expected area of the square.

MH 2500 Probability and introduction to statistics

Tutorial 5 Week 6. Chapters 2.2-2.3, 4.1. Solutions

1. The length of a human pregnancy can be considered to be a normally distributed random variable with a mean of 266 days and a standard deviation of 16 days. If we let

$$X = \text{The length of a human pregnancy}$$

it follows from the assumptions that $X \sim N(266, 16^2)$ where the unit is days.

(a) Determine the probability that the length of the pregnancy for a randomly chosen pregnant woman will be less than 240 days, i.e. that it, roughly speaking, will last less than 8 months. In order to be able to use the normal table we first have to standardize.

$$\Pr(X < 240) = \Pr\left(Z < \frac{240 - 266}{16} = -1.625\right)$$

Due to the fact that the standard normal distribution is *symmetric about zero*, it follows that

$$\Pr(Z < -1.625) = \Pr(Z > 1.625)$$

Due to the construction of Table 2 of Appendix B we make use of the complement rule to find that

$$\Pr(Z > 1.625) = 1 - \Pr(Z < 1.625) = 1 - \Phi(1.625) \approx \mathbf{0.052}$$

We interpret this as on average every twentieth pregnancy lasts less than 240 days.

(b) Determine the third quartile, q_3. Interpret this value. According to Table 2 of Appendix B

$$z_{0.75} \approx 0.675$$

which means that

$$q_3 = x_{0.75} = 266 + 0.675 \cdot 16 \approx \mathbf{276.8}$$

We interpret this as, on average, every fourth pregnancy lasts longer than 277 days.

(c) Suppose that we study 5 randomly chosen pregnant women. Determine the probability that at least two of them will be "late", i.e. have a pregnancy that lasts longer than expected. Let

$$Y = \text{The number of women (in this sample) that will be "late"}$$

A pregnant woman is either "late" or is not "late", i.e. if we consider a woman who is "late" as a "success" we have a situation where we draw balls from an urn that only contains "successes" and "failures". If we assume that pregnant women are "late" independently of each other, the draw is conducted *with* replacement. Since the normal distribution is symmetric about its mean, the probability that a randomly chosen pregnant woman will be "late" is 0.5. Finally, since our random variable is counting the number of "successes" among the drawn balls it is clear that $Y \sim Bi\,(5, 0.5)$. And so

$$\Pr(Y \geq 2) = 1 - \Pr(Y \leq 1) = 1 - \left[\binom{5}{0}0.5^5 + \binom{5}{1}0.5^5\right] = \mathbf{0.8125}$$

2. (**2.62**) A particle of mass m has a random velocity, V, which is normally distributed with parameters $\mu = 0$ and σ. Find the pdf of the kinetic energy, $E = \frac{1}{2}mV^2$.
Remark. This problem is similar to the situation discussed in Example C (p 61) in the textbook, and the technique used to solve it is practically identical.
Since the transformation is *not* strictly monotone, we use the cdf-method.

Step 1. Expressing the cdf of E in terms of the cdf of V.

$$\begin{aligned} F_E(x) &= \Pr(E \leq x) = \Pr\left(\frac{1}{2}mV^2 \leq x\right) = \Pr\left(-\sqrt{\frac{2x}{m}} \leq V \leq \sqrt{\frac{2x}{m}}\right) = \\ &= F_V\left(\sqrt{\frac{2x}{m}}\right) - F_V\left(-\sqrt{\frac{2x}{m}}\right), \quad x > 0 \end{aligned}$$

Step 2. Differentiating the cdf of E (now expressed in terms of the cdf of V) with respect to x.

$$\begin{aligned} f_E(x) &= \frac{f_V\left(\sqrt{\frac{2x}{m}}\right) + f_V\left(-\sqrt{\frac{2x}{m}}\right)}{\sqrt{2mx}} = \sqrt{\frac{2}{mx}} f_V\left(\sqrt{\frac{2x}{m}}\right) = \\ &= \sqrt{\frac{2}{mx}} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\left(\sqrt{\frac{2x}{m}}\right)^2\right\} = \\ &= \frac{1}{\sigma\sqrt{\pi m x}} \exp\left\{-\frac{x}{m\sigma^2}\right\}, \quad x > 0 \end{aligned}$$

At this point it is possible to further rewrite the expression, but we have to be familiar with the fact that $\sqrt{\pi} = \Gamma(1/2)$.

$$\begin{aligned} f_E(x) &= \frac{1}{\sigma\sqrt{\pi m x}} \exp\left\{-\frac{x}{m\sigma^2}\right\} = \\ &= \frac{1}{\Gamma\left(\frac{1}{2}\right)(m\sigma^2)^{1/2}} x^{\frac{1}{2}-1} \exp\left\{-\frac{x}{m\sigma^2}\right\}, \quad x > 0 \end{aligned}$$

and it is clear that $E \sim Ga\left(\frac{1}{2}, \frac{1}{m\sigma^2}\right)$.

3. (**4.20**) Let $X \sim Po(\lambda)$. Find

$$E\left(\frac{1}{X+1}\right)$$

First we just wright out the expectation as a sum and follow the "standard procedure" and merge k into the pmf of X.

$$E\left(\frac{1}{X+1}\right) = \sum_{k=0}^{\infty} \frac{1}{k+1} \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k+1)!}$$

Next we use multiplicative constants to manipulate the summand so that it becomes a pmf of a (possibly different) poisson random variable

$$E\left(\frac{1}{X+1}\right) = \sum_{k=0}^{\infty} \frac{1}{k+1} \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k+1)!} = \frac{1}{\lambda}\sum_{k=0}^{\infty} \frac{\lambda^{k+1} e^{-\lambda}}{(k+1)!} = \frac{1}{\lambda}\sum_{j=1}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!}$$

Finally, we realize that the sum just represents $\Pr(X \geq 1)$ for a $Po(\lambda)$, and so

$$E\left(\frac{1}{X+1}\right) = \frac{1}{\lambda}\Pr(X \geq 1) = \frac{1}{\lambda}(1 - \Pr(X = 0)) = \frac{1 - e^{-\lambda}}{\lambda}$$

4. (**4.23**) A random square has a side length that is an exponential random variable. Find the expected area of the square. If we let X represent the side length of the square it is clear that $X \sim Exp(\lambda)$. The expected area of the square is

$$E\left(X^2\right) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx$$

If we inspect the integrand it looks similar to the pdf of a $Ga(3, \lambda)$. Since the area of integration is identical to the domain of this pdf we use multiplicative constants to manipulate the integrand so that it becomes this particular pdf.

$$E\left(X^2\right) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}\int_0^{\infty} \frac{\lambda^3}{2} x^2 e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

MH 2500 Probability and introduction to statistics

Tutorial 6 Week 7. Chapter 2, Section 4.1. Problems

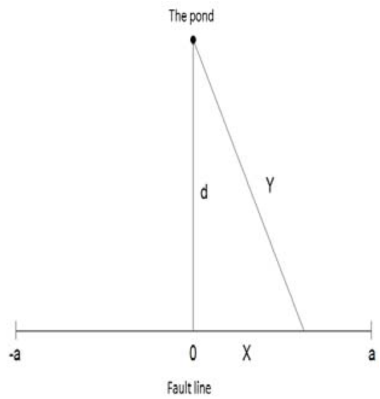
1. Let X be a random variable with mean μ and standard deviation σ. Determine $\Pr(X > \mu + \sigma)$ for the following distributions:

   (a) $X \sim N(\mu = 7,\ \sigma^2 = 0.16)$.

   (b) $X \sim Bi(n = 5,\ p = 0.4)$.

   (c) $X \sim Hyp(n = 5,\ r = 5,\ N = 25)$.

2. The epicentre of an earthquake ends up, with uniform probability distribution, along a fault line of length $2a$ kilometres. A pond is located d kilometres from the midpoint of this fault line. Let X represent the location of the epicentre along the fault line, and let Y be the distance from the pond to the epicentre.



   Determine the pdf of Y.

3. (**4.8**) Show that if X is a discrete random variable, taking values on the positive integers, then $E(X) = \sum_{k=1}^{\infty} \Pr(X \geq k)$. Apply this result to find the expected value of a geometric random variable.

MH 2500 Probability and introduction to statistics

Tutorial 6 Week 7. Chapter 2, Section 4.1. Solutions

1. Let X be a random variable with mean μ and standard deviation σ. Determine $\Pr(X > \mu + \sigma)$ for the following distributions:

   (a) $X \sim N(\mu = 7,\ \sigma^2 = 0.16)$. Since

   $$\mu + \sigma = 7 + 0.4 = 7.4$$

   standardization yields

   $$\Pr(X > 7.4) = \Pr\left(Z > \frac{7.4 - 7}{0.4} = 1\right) = 1 - 0.8413 = \mathbf{0.1587}$$

   just as for *all* normal distributions.

   (b) $X \sim Bi(n = 5,\ p = 0.4)$. Now

   $$\begin{aligned} \mu &= np = 5 \cdot 0.4 = 2 \\ \sigma &= \sqrt{np(1-p)} = \sqrt{5 \cdot 0.4 \cdot 0.6} = 1.0954 \end{aligned}$$

   This means that we want to determine

   $$\Pr(X > 2 + 1.0954 = 3.0954)$$

   which, since X is discrete, is the same as

   $$\Pr(X \geq 4) = \Pr(X = 4) + \Pr(X = 5) = \tbinom{5}{4} \cdot 0.4^4 \cdot 0.6^1 + \tbinom{5}{5} \cdot 0.4^5 \cdot 0.6^0 = \mathbf{0.087}$$

   (c) $X \sim Hyp(n = 5,\ r = 5,\ N = 25)$. Now

   $$\begin{aligned} \mu &= \frac{nr}{N} = \frac{5 \cdot 5}{25} = 1 \\ \sigma &= \sqrt{\frac{nr(N-r)(N-n)}{N^2(N-1)}} = \sqrt{\frac{5 \cdot 5 \cdot (25-5) \cdot (25-5)}{25^2 \cdot (25-1)}} = \sqrt{\frac{2}{3}} \approx 0.816 \end{aligned}$$

   This means that we want to determine

   $$\Pr(X > 1 + 0.816 = 1.816)$$

   which, since X is discrete, is the same as

   $$\Pr(X \geq 2) = 1 - \Pr(X \leq 1) = 1 - \frac{\binom{5}{0}\binom{20}{5} + \binom{5}{1}\binom{20}{4}}{\binom{25}{5}} = \mathbf{0.2522}$$

2. The epicentre of an earthquake ends up, with uniform probability distribution, along a fault line of length $2a$ kilometres. A pond is located d kilometres from the midpoint of this fault line. Let X represent the location of the epicentre along the fault line, and let Y be the distance from the pond to the epicentre. Determine the pdf of Y. If we let $x = 0$ represent the midpoint of the fault line, it is clear that $X \sim U(-a, a)$, that is,

$$f_X(x) = \frac{1}{2a}, \quad -a \le x \le a$$

Furthermore, we have that

$$Y = \sqrt{X^2 + d^2}$$

Applying the cdf-method, we first express the cdf of Y in terms of the cdf of Y.

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) = \Pr\left(\sqrt{X^2 + d^2} \le y\right) = \Pr\left(X^2 \le y^2 - d^2\right) = \\ &= \Pr\left(-\sqrt{y^2 - d^2} \le X \le \sqrt{y^2 - d^2}\right) = F_X\left(\sqrt{y^2 - d^2}\right) - F_X\left(-\sqrt{y^2 - d^2}\right) \end{aligned}$$

Differentiating with respect to y yields

$$\begin{aligned} f_Y(y) &= f_X\left(\sqrt{y^2 - d^2}\right) \cdot \frac{y}{\sqrt{y^2 - d^2}} + f_X\left(-\sqrt{y^2 - d^2}\right) \cdot \frac{y}{\sqrt{y^2 - d^2}} = \\ &= \frac{y}{a\sqrt{y^2 - d^2}}, \quad d \le y \le \sqrt{d^2 + a^2} \end{aligned}$$

3. (**4.8**) Show that if X is a discrete random variable, taking values on the positive integers, then $E(X) = \sum_{k=1}^{\infty} \Pr(X \geq k)$. Apply this result to find the expected value of a geometric random variable. Consider a discrete random variable with range $k = 1, 2, 3, \ldots$. An argument goes as follows. According to the definition of the expected value of X,

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k)$$

Studying this expression, it is clear that we can interpret $E(X)$ as a sum in which the probability associated with a certain positvie integer k, i.e. $\Pr(X = k)$, is counted k times. If we, on the other hand, study the sum

$$\sum_{k=1}^{\infty} \Pr(X \geq k)$$

follows that $\Pr(X = k)$ is included exactly k times, one time each in the terms $\Pr(X \geq 1), \Pr(X \geq 2), \ldots, \Pr(X \geq k)$. We can visualize the situation as follows

	$\Pr(X \geq 1)$	$\Pr(X \geq 2)$	$\Pr(X \geq 3)$	$\Pr(X \geq 4)$	$\cdots$
$1 \cdot \Pr(X = 1)$	$\Pr(X = 1)$				$\cdots$
$2 \cdot \Pr(X = 2)$	$\Pr(X = 2)$	$\Pr(X = 2)$			$\cdots$
$3 \cdot \Pr(X = 3)$	$\Pr(X = 3)$	$\Pr(X = 3)$	$\Pr(X = 3)$		$\cdots$
$4 \cdot \Pr(X = 4)$	$\Pr(X = 4)$	$\Pr(X = 4)$	$\Pr(X = 4)$	$\Pr(X = 4)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Summing the terms row-wise or column-wise should not make any difference. A formal calculation is

$$\begin{aligned}\sum_{k=1}^{\infty} \Pr(X \geq k) &= \sum_{k=1}^{\infty} \sum_{i=k}^{\infty} \Pr(X = i) = \sum_{i=1}^{\infty} \sum_{k=1}^{i} \Pr(X = i) = \\ &= \sum_{i=1}^{\infty} \Pr(X = i) \sum_{k=1}^{i} 1 = \sum_{i=1}^{\infty} i \Pr(X = i) = E(X)\end{aligned}$$

where we att one step change the order of summation (which is equivalent to switch from column-wise to row-wise calculation).

Now let $X \sim Ge\,(p)$. Since

$$\Pr\left(X = k\right) = \left(1 - p\right)^{k-1} p, \quad k = 1, 2, \ldots$$

it follows that

$$\Pr\left(X \geq k\right) = \sum_{i=k}^{\infty} \left(1 - p\right)^{i-1} p = p \left(\sum_{i=1}^{\infty} \left(1 - p\right)^{i-1} - \sum_{i=1}^{k-1} \left(1 - p\right)^{i-1} \right)$$

Next we use results for both geometric series and geometric sums to conclude that

$$\Pr\left(X \geq k\right) = p \left(\frac{1}{1 - (1 - p)} - \frac{1 - (1 - p)^{k-1}}{1 - (1 - p)} \right) = \left(1 - p\right)^{k-1}$$

and so, again using results for geometric series,

$$E\left(X\right) = \sum_{k=1}^{\infty} \Pr\left(X \geq k\right) = \sum_{k=1}^{\infty} \left(1 - p\right)^{k-1} = \frac{1}{1 - (1 - p)} = \frac{1}{p}$$

MH 2500 Probability and introduction to statistics

Tutorial 7 Week 8. Chapters 2.2-2.3, 4.1. Problems

1. The Normal Distribution.

   (a) (**2.56**) If $X \sim N(\mu, \sigma^2)$, show that $\Pr(|X-\mu| \leq 0.675\sigma) = 0.5$.

   (b) (**2.55**) If $X \sim N(\mu, \sigma^2)$, find the value of c in terms of σ such that $\Pr(\mu - c \leq X \leq \mu + c) = 0.95$.

2. Let $X \sim Beta\,(a, b)$. Find the mean and the variance of X.

3. Consider a probability distribution with mean μ and variance σ^2. The most common way to define the *skewness* of a probability distribution is to use the *third standardized moment*, i.e.
$$E\left(\frac{X-\mu}{\sigma}\right)^3.$$
It turns out that this expression can be rewritten as
$$E\left(\frac{X-\mu}{\sigma}\right)^3 = \frac{E(X^3) - 3\mu\sigma^2 - \mu^3}{\sigma^3} = \frac{E(X^3) - 3\mu\sigma^2 - \mu^3}{(\sigma^2)^{3/2}}.$$
Show that the skewness of the gamma distribution, $Ga(\alpha, \lambda)$, is $2/\sqrt{\alpha}$.
Remark. We will start looking at the concept of moments during the Thursday lecture.

4. (**4.26**) A stick of unit length is broken into two pieces. Find the expected ratio of the length of the longer piece to the length of the shorter piece.
Hint. Let X be the position of the breakage. We don't have much information about the distribution of X. So which distribution is the most natural to use? Now let Y be the ratio of the length of the longer piece to the length of the shorter piece. Try to find the cdf of Y.

MH 2500 Probability and introduction to statistics

Tutorial 7 Week 8. Chapters 2.2-2.3, 4.1. Solutions

1. The Normal Distribution.

   (a) (**2.56**) If $X \sim N(\mu, \sigma^2)$, show that $\Pr(|X-\mu| \leq 0.675\sigma) = 0.5$. We are to show that for any normal distribution, half of the probability mass is located within 0.675 standard deviations from the mean. We have that

$$\begin{aligned}
\Pr(|X-\mu| \leq 0.675\sigma) &= \Pr(-0.675\sigma \leq X-\mu \leq 0.675\sigma) = \\
&= \Pr\left(-0.675 \leq \frac{X-\mu}{\sigma} \leq 0.675\right) = \\
&= \Pr(-0.675 \leq Z \leq 0.675) = \\
&= \Pr(Z \leq 0.675) - \Pr(Z < -0.675) = \\
&= \Pr(Z \leq 0.675) - [1 - \Pr(Z \leq 0.675)] = \\
&= 2\Pr(Z \leq 0.675) - 1 = 2 \cdot 0.75 - 1 = 0.5
\end{aligned}$$

   (b) (**2.55**) If $X \sim N(\mu, \sigma^2)$, find the value of c in terms of σ such that $\Pr(\mu - c \leq X \leq \mu + c) = 0.95$. Using the same approach as in the previous problem we get that

$$\begin{aligned}
0.95 &= \Pr(\mu - c \leq X \leq \mu + c) = \Pr\left(-\frac{c}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{c}{\sigma}\right) = \\
&= \Pr\left(-\frac{c}{\sigma} \leq Z \leq \frac{c}{\sigma}\right) = \Pr\left(Z \leq \frac{c}{\sigma}\right) - \left[1 - \Pr\left(Z \leq \frac{c}{\sigma}\right)\right] = \\
&= 2\Pr\left(Z \leq \frac{c}{\sigma}\right) - 1
\end{aligned}$$

   meaning that

$$\Pr\left(Z \leq \frac{c}{\sigma}\right) = \frac{0.95+1}{2} = 0.975$$

   According to the table of $N(0,1)$, this implies that $c/\sigma = 1.96$ and therefore $c = 1.96\sigma$. Hence, for any normal distribution 95% of the probability mass is located within 1.96 standard deviations from the mean.

2. Let $X \sim Beta\,(a,b)$. Find the mean and the variance of X. By definition of expectation,

$$\mu = E\,(X) = \int_0^1 x \cdot \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} x^{a-1}\,(1-x)^{b-1}\,dx = \int_0^1 \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} x^{a}\,(1-x)^{b-1}\,dx.$$

The integrand look similar to the pdf of $Beta\,(a+1,b)$, and so

$$\mu = \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} \cdot \frac{\Gamma\,(a+1)\,\Gamma\,(b)}{\Gamma\,(a+1+b)} \int_0^1 \frac{\Gamma\,(a+1+b)}{\Gamma\,(a+1)\,\Gamma\,(b)} x^{a}\,(1-x)^{b-1}\,dx = \frac{a}{a+b}.$$

In a similar way it follows that

$$\begin{aligned} E\left(X^2\right) &= \int_0^1 x^2 \cdot \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} x^{a-1}\,(1-x)^{b-1}\,dx = \int_0^1 \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} x^{a+1}\,(1-x)^{b-1}\,dx = \\ &= \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} \cdot \frac{\Gamma\,(a+2)\,\Gamma\,(b)}{\Gamma\,(a+2+b)} \int_0^1 \frac{\Gamma\,(a+2+b)}{\Gamma\,(a+2)\,\Gamma\,(b)} x^{a}\,(1-x)^{b-1}\,dx = \\ &= \frac{a\,(a+1)}{(a+b)\,(a+b+1)}, \end{aligned}$$

where we in the last step use the fact that the integrand is the pdf of $Beta\,(a+2,b)$. Finally, according to the calculation formula for the variance,

$$\begin{aligned} Var\,(X) &= E\left(X^2\right) - \mu^2 = \frac{a\,(a+1)}{(a+b)\,(a+b+1)} - \left(\frac{a}{a+b}\right)^2 = \\ &= \frac{a\,(a+1)\,(a+b) - a^2\,(a+b+1)}{(a+b)^2\,(a+b+1)} = \frac{ab}{(a+b)^2\,(a+b+1)}. \end{aligned}$$

3. Consider a probability distribution with mean μ and variance σ^2. The most common way to define the *skewness* of a probability distribution is to use the *third standardized moment*, i.e.

$$E\left(\frac{X-\mu}{\sigma}\right)^3.$$

It turns out that this expression can be rewritten as

$$E\left(\frac{X-\mu}{\sigma}\right)^3 = \frac{E\left(X^3\right)-3\mu\sigma^2-\mu^3}{\sigma^3} = \frac{E\left(X^3\right)-3\mu\sigma^2-\mu^3}{\left(\sigma^2\right)^{3/2}}.$$

Show that the skewness of the gamma distribution, $Ga(\alpha,\lambda)$, is $2/\sqrt{\alpha}$. From our list of common distributions we find that $\mu=\alpha/\lambda$ and $\sigma^2=\alpha/\lambda^2$. Now

$$E\left(X^3\right) = \int_0^\infty x^3\cdot\frac{\lambda^\alpha}{\Gamma\left(\alpha\right)}x^{\alpha-1}e^{-\lambda x}dx = \int_0^\infty \frac{\lambda^\alpha}{\Gamma\left(\alpha\right)}x^{\alpha+2}e^{-\lambda x}dx.$$

The integrand look similar to the pdf of $Ga(\alpha+3,\lambda)$, and so

$$E\left(X^3\right) = \frac{\lambda^\alpha}{\Gamma\left(\alpha\right)}\cdot\frac{\Gamma\left(\alpha+3\right)}{\lambda^{\alpha+3}}\int_0^\infty\frac{\lambda^{\alpha+3}}{\Gamma\left(\alpha+3\right)}x^{\alpha+2}e^{-\lambda x}dx = \frac{\alpha\left(\alpha+1\right)\left(\alpha+2\right)}{\lambda^3}.$$

Finally,

$$\begin{aligned} E\left(\frac{X-\mu}{\sigma}\right)^3 &= \frac{E\left(X^3\right)-3\mu\sigma^2-\mu^3}{\left(\sigma^2\right)^{3/2}} = \frac{\frac{\alpha(\alpha+1)(\alpha+2)}{\lambda^3}-3\cdot\frac{\alpha}{\lambda}\cdot\frac{\alpha}{\lambda^2}-\left(\frac{\alpha}{\lambda}\right)^3}{\frac{\alpha^{3/2}}{\lambda^3}} = \\ &= \frac{\alpha\left(\alpha+1\right)\left(\alpha+2\right)-3\alpha^2-\alpha^3}{\alpha\sqrt{\alpha}} = \frac{2}{\sqrt{\alpha}}. \end{aligned}$$

4. (**4.26**) A stick of unit length is broken into two pieces. Find the expected ratio of the length of the longer piece to the length of the shorter piece. We first let

$$X = \text{The position of the breakage}$$

The only information we have is that the stick is of unit length. We therefore assume that the position of the breakage is uniform, i.e. $X \sim U(0,1)$. Now let

$$Y = \text{The ratio of the length of the longer piece to the length of the shorter piece}$$

and it is clear that the range of Y (i.e. the domain of f_Y) is $y \geq 1$. Furtermore, the cdf of Y is given by

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr\left(\frac{1}{y+1} \leq X \leq \frac{y}{y+1}\right) = \\ &= \frac{y-1}{y+1} = 1 - \frac{2}{y+1}, \quad y \geq 1, \end{aligned}$$

which leads to

$$f_Y(y) = \frac{2}{(y+1)^2}, \quad y \geq 1,$$

and so

$$E(Y) = \int_1^{\infty} \frac{2y}{(y+1)^2} dy = \infty.$$

We therefore conclude that $E(Y)$ doesn't exist.

MH 2500 Probability and introduction to statistics

Tutorial 8 Week 9. Chapters 3.1-3.4, 4.5. Problems

1. (**4.91**) Let $X \sim Exp(\lambda)$ and c be a positive real constant. Determine the moment generating function (mgf) of X. Use the mgf of X to show that $Y = cX$ is also an exponentially distributed random variable.

2. Let $X \sim Bin(n, p)$.

   (a) Use moment generating functions to show that $Y = n - X$ is also a binomial random variable.

   (b) Present an argument that (informally) proves that $Y = n - X$ is also a binomial random variable.

3. The joint pdf of X and Y is given by

$$f_{X,Y}(x, y) = 6, \quad 0 \le x \le 1, \quad x^2 \le y \le x.$$

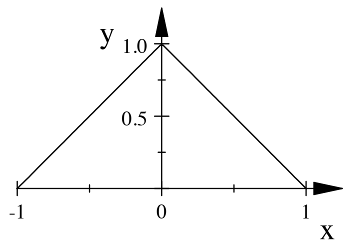
   (a) Visualize the domain of $f_{X,Y}(x, y)$.

   (b) Determine the marginal distribution of X, i.e. find $f_X(x)$.

   (c) Determine the mean and the variance of X.

   (d) Are X and Y independent? Motivate.

4. Suppose that X and Y are uniformly distributed over the triangle shown in the figure below.



   Determine the marginal distributions of X and Y, i.e. find $f_X(x)$ and $f_Y(y)$.

MH 2500 Probability and introduction to statistics

Tutorial 8 Week 9. Chapters 3.1-3.4, 4.5. Solutions

1. (**4.91**) Let $X \sim Exp(\lambda)$ and c be a positive real constant. Determine the moment generating function (mgf) of X. Use the mgf of X to show that $Y = cX$ is also an exponentially distributed random variable. First we uses the standard technique to find the mgf of X.

$$M(t) = E\left(e^{tX}\right) = \int_0^\infty e^{tx} \cdot \lambda e^{-\lambda x} dx = \int_0^\infty \lambda e^{-\lambda x + tx} dx = \int_0^\infty \lambda e^{-(\lambda - t)x} dx.$$

The integrand look similar to the pdf of $Exp(\lambda - t)$ so

$$M(t) = \frac{\lambda}{\lambda - t} \int_0^\infty (\lambda - t) e^{-(\lambda - t)x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda,$$

where the restriction on t follows from the fact that the parameter of the exponential distribution, i.e. the intensity of the associated poisson process, must be positive. We make the observation that $Y = cX$ is a linear transformation on the form $Y = a + bX$ where $a = 0$ and $b = c$, which implies that

$$M_Y(t) = e^0 M_X(ct) = \frac{\lambda}{\lambda - ct} = \frac{\lambda/c}{\lambda/c - t}, \quad t < \lambda/c.$$

We recognize this as the mgf of $Exp(\lambda/c)$ and therefore, since the mgf is unique, we conclude that $Y \sim Exp(\lambda/c)$.

2. Let $X \sim Bin(n, p)$.

   (a) Use moment generating functions to show that $Y = n - X$ is also a binomial random variable.Let $q = 1 - p$. We make the observation that $Y = n - X$ is a linear transformation on the form $Y = a + bX$ where $a = n$ and $b = -1$, which implies that

$$m_Y(t) = e^{tn} m_Y(-t) = e^{tn}\left(pe^{-t} + q\right)^n = \left[e^t\left(pe^{-t} + q\right)\right]^n = \left(qe^t + p\right)^n$$

   We recognize this as the mgf of $Bin(n, q)$ and therefore, since the mgf is unique, we conclude that $Y \sim Bin(n, q)$. If you are not familiar with the result concerning the mgf of a linear transformation, it is almost as easy to do it the "hard way".

$$m_Y(t) = E\left(e^{tX}\right) = E\left(e^{t(n-Y)}\right) = E\left(e^{tn}e^{-tY}\right) = e^{tn}E\left(e^{-tY}\right) = e^{tn}m_Y(-t)$$

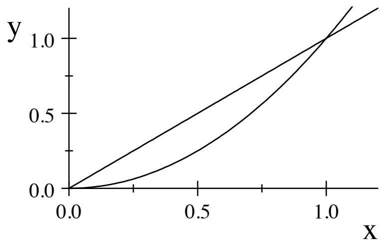
   just as above.

(b) Present an argument that (informally) proves that $Y = n - X$ is also a binomial random variable. Since $X \sim Bin\,(n, p)$, we have a situation where we are to draw, with replacement, n balls from an urn where the constant probability of success is p. X counts the number of successes among the drawn balls. Therefore $Y = n - X$ counts the number of failures among the drawn balls. Hence, we have an almost identical situation where the only difference is that we count failures instead of successes, and since the constant probability of failure is $q = 1 - p$ it must hold that $Y \sim Bin\,(n, q)$.

3. The joint pdf of X and Y is given by

$$f_{X,Y}\,(x, y) = 6, \quad 0 \leq x \leq 1, \quad x^2 \leq y \leq x.$$

(a) Visualize the domain of $f_{X,Y}\,(x, y)$.



(b) Determine the marginal distribution of X, i.e. find $f_X\,(x)$. We find the marginal distribution of X via

$$f_X\,(x) = 6 \int_{x^2}^{x} dy = 6\left(x - x^2\right) = 6x\,(1 - x)\,, \quad 0 \leq x \leq 1.$$

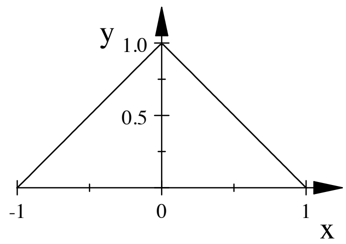
which means that $X \sim Beta\,(2, 2)$.

(c) Determine the mean and the variance of X. Since $X \sim Beta\,(2, 2)$ it follows from the information that we have about the beta distribution that

$$\begin{aligned} E\,(X) &= \frac{2}{2+2} = \mathbf{\frac{1}{2}} \\ Var\,(X) &= \frac{2 \cdot 2}{(2+2)^2\,(2+2+1)} = \mathbf{\frac{1}{20}} \end{aligned}$$

(d) Are X and Y independent? Motivate. Since the domain of $f_{X,Y}\,(x, y)$ is not rectangular X and Y cannot be independent.

4. Suppose that X and Y are uniformly distributed over the triangle shown in the figure below.



Determine the marginal distributions of X and Y, i.e. find $f_X(x)$ and $f_Y(y)$. Since the distribution is uniform and the area of the triangle is 1, it follows that the joint pdf of X and Y is given by

$$f_{X,Y}(x,y) = 1,\ -1 \leq x \leq 1,\ 0 \leq y \leq 1 - |x|,$$

where we have made the important observation that for a given value of x, such that $-1 \leq x \leq 1$, it follows that $0 \leq y \leq 1 - |x|$. This observation helps us to show that the marginal pdf of X is given by

$$f_X(x) = \int_0^{1-|x|} dy = 1 - |x|,\ -1 \leq x \leq 1.$$

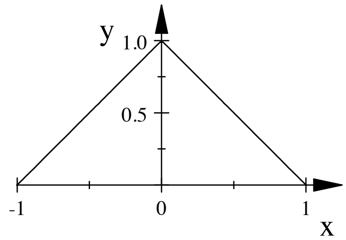
In order to find the marginal pdf of y, we make the observation that for a given value of y, such that $0 \leq y \leq 1$, it follows that $y - 1 \leq x \leq 1 - y$. Hence,

$$f_Y(y) = \int_{y-1}^{1-y} dx = 1 - y - (y - 1) = 2(1 - y),\ 0 \leq y \leq 1.$$

MH 2500 Probability and introduction to statistics

Tutorial 9 Week 10. Chapters 3.5-3.6.1, 4.5. Problems

1. Suppose that X and Y are uniformly distributed over the triangle shown in the figure below.



   (a) Show that $Y \mid X = x$ is a uniformly distributed random variable. Use this fact to quickly determine $\Pr(Y > 1/2 \mid X = 1/4)$.

   (b) Let $W = Y - X$. Determine the probability distribution of W, i.e. find $f_W(w)$. *Hint.* First find $F_W(w)$. See lecture notes.

2. (**Rice 3.45**) For a Poisson distribution, suppose that events are independently labeled A and B with probabilities $p_A + p_B = 1$. If the parameter of the Poisson distribution is λ, show that the number of events labeled A follows a Poisson distribution with parameter $p_A\lambda$.
   Hint. This is a hierarchic model. For a more detailed hint, see the answers to the textbook problems provided on NTU Learn.

3. (**Rice 4.87**) Under what conditions is the sum of independent negative binomial random variables also negative binomial?

4. Let X and Y be independent random variables whose moment generating functions are given by

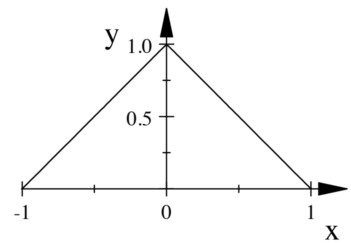
$$M_X(t) = \exp\left\{2e^t - 2\right\}, \text{ and } M_Y(t) = \left(\frac{3}{4}e^t + \frac{1}{4}\right)^{10},$$

   respectively. Compute $\Pr(XY = 0)$.

MH 2500 Probability and introduction to statistics

Tutorial 9 Week 10. Chapters 3.5-3.6.1, 4.5. Solutions

1. Suppose that X and Y are uniformly distributed over the triangle shown in the figure below.



(a) Show that $Y \mid X = x$ is a uniformly distributed random variable. Use this fact to quickly determine $\Pr(Y > 1/2 \mid X = 1/4)$. From the previous Tutorial we know that

$$f_{X,Y}(x, y) = 1, \ -1 \leq x \leq 1, \quad 0 \leq y \leq 1 - |x|,$$

and

$$f_X(x) = \int_0^{1-|x|} dy = 1 - |x|, \quad -1 \leq x \leq 1.$$

By the definition of conditional pdf it therefore follows that

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{1}{1 - |x|}, \quad 0 \leq y \leq 1 - |x|.$$

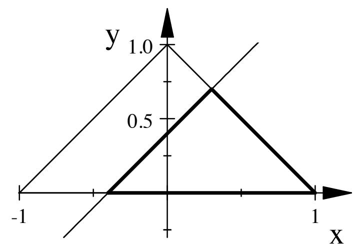
Hence, we conclude that $Y \mid X = x \sim U(0, 1 - |x|)$. Therefore $Y \mid X = \frac{1}{4} \sim U\left(0, \frac{3}{4}\right)$ and, since the distribution is uniform, it finally follows that

$$\Pr(Y > 1/2 \mid X = 1/4) = \frac{3/4 - 1/2}{3/4} = \frac{1/4}{3/4} = \mathbf{\frac{1}{3}}$$

(b) Let $W = Y - X$. Determine the probability distribution of W, i.e. find $f_W(w)$. *Hint.* First find $F_W(w)$. See lecture notes. Following the steps presented in the lecture notes, we first find a general expression for the cdf of W, i.e.

$$F_W(w) = \Pr(W \leq w) = \Pr(Y - X \leq w) = \Pr(Y \leq X + w).$$

So in order to find $F_W(w)$ we have to integrate $f_{X,Y}(x, y)$ over the part of the domain where $y \leq x + w$. To get everything right, a good practice is to describe the domain and area of integration graphically, i.e.



where we for an arbitrarily chosen w have included the line $y = x + w$. It is clear that the area of integration is the triangle in the lower right corner of the domain, and since $f_{X,Y}(x, y) = 1$ we have that $F_W(w)$ equals the *area* of this triangle. Since the length of the base of the triangle is $1 + w$ and the height is $(1 + w)/2$, it follows that the area of the triangle is $(1 + w)^2/4$. Furthermore, changing the value of w will result in a parallel translation of the line in the graph, and therefore it is clear that the relevant values of w ranges from -1 to 1. We thus conclude that

$$F_W(w) = \frac{(1 + w)^2}{4}, \quad -1 \leq w \leq 1.$$

Finally, differentiating with respect to w yields

$$f_W(w) = \frac{(1 + w)}{2}, \quad -1 \leq w \leq 1.$$

Remark. A much quicker solution would be possible if we had access to the general formula for how to deal with a difference of two random variables. See Problem 3.50 in the textbook.

2. (**Rice 3.45**) For a Poisson distribution, suppose that events are independently labeled A and B with probabilities $p_A + p_B = 1$. If the parameter of the Poisson distribution is λ, show that the number of events labeled A follows a Poisson distribution with parameter $p_A\lambda$. If we let N represent the number of events that occur on an interval of unit length, then, by definition, $N \sim Po(\lambda)$. Furthermore, let X represent the number of events that have been labeled A. By the assumptions, it then holds that $X \mid N = n$ är $Bi(n, p_A)$. We are looking for the marginal (or unconditioned) distribution of X, i.e. the marginal pmf p_X. It follows from the Law of total probability for pmfs that

$$\begin{aligned} p_X(x) &= \sum_{n=0}^{\infty} p_{X,N}(x,n) = \sum_{n=0}^{\infty} p_{X|N}(x \mid n)\, p_N(n) = \sum_{n=x}^{\infty} \binom{n}{x} p_A^x p_B^{n-x} \frac{\lambda^n e^{-\lambda}}{n!} = \\ &= \frac{(\lambda p_A)^x e^{-\lambda}}{x!} \sum_{n=x}^{\infty} \frac{(\lambda p_B)^{n-x}}{(n-x)!} \overset{j=n-x}{=} \frac{(\lambda p_A)^x e^{-\lambda} e^{\lambda p_B}}{x!} \sum_{j=0}^{\infty} \frac{(\lambda p_B)^j e^{-\lambda p_B}}{j!} = \\ &= \frac{(\lambda p_A)^x e^{-\lambda(1-p_B)}}{x!} = \frac{(\lambda p_A)^x e^{-\lambda p_A}}{k!}, \quad x = 0, 1, \ldots \end{aligned}$$

and we therefore conclude that $X \sim Po(\lambda p_A)$.

3. (**Rice 4.87**) Under what conditions is the sum of independent negative binomial random variables also negative binomial? Let $X_1, X_2, \ldots, X_n$ be a collection of mutually independent random variables such that $X_i \sim NegBin(r_i, p_i)$ for $i = 1, 2, \ldots, n$. Furthermore, let $Y = \sum_{i=1}^n X_i$. According to (a generalized version of) Property D in Section 4.5 of the textbook (also proven during lecture),

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} \left(\frac{p_i e^t}{1 - (1 - p_i) e^t} \right)^{r_i}$$

In order for us to be able to further simplify this expression, a requirement is that $p_i = p$ for all i which in that case means that

$$M_Y(t) = \prod_{i=1}^{n} \left(\frac{p e^t}{1 - (1 - p) e^t} \right)^{r_i} = \left(\frac{p e^t}{1 - (1 - p) e^t} \right)^{\sum_{i=1}^n r_i}$$

from which we conclude that $Y \sim NegBin(\sum_{i=1}^n r_i, p)$.

4. Let X and Y be independent random variables whose moment generating functions are given by

$$M_X(t) = \exp\left\{2e^t - 2\right\}, \text{ and } M_Y(t) = \left(\frac{3}{4}e^t + \frac{1}{4}\right)^{10},$$

respectively. Compute $\Pr(XY = 0)$. Checking our list of mgfs for the common probability distributions, we observe that $X \sim Po(2)$ and that $Y \sim Bi(10, 3/4)$. In order for the product XY to equal zero, *at least* one of X and Y must equal zero. Let A be the event that $X = 0$, and B the event that $Y = 0$. First,

$$\begin{aligned}
\Pr(A) &= \Pr(X = 0) = \frac{2^0 e^{-2}}{0!} = e^{-2} \\
\Pr(B) &= \Pr(Y = 0) = \tbinom{10}{0}\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)^{10} = \frac{1}{4^{10}}
\end{aligned}$$

Now it follows by the Addition Law and the fact that A and B are independent events that

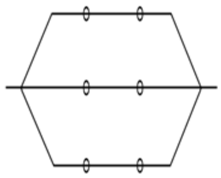
$$\begin{aligned}
\Pr(XY = 0) &= \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = \\
&= \Pr(A) + \Pr(B) - \Pr(A) \cdot \Pr(B) = \\
&= e^{-2} + \frac{1}{4^{10}} - \frac{e^{-2}}{4^{10}} \approx \mathbf{0.135}
\end{aligned}$$

MH 2500 Probability and introduction to statistics

Tutorial 10 Week 11. Sections 3.6.1, 3.7, 4.3, 4.5. Problems

1. In a missile-testing program, one random variable of interest is the distance between the point at which the missile lands and the center of the target at which the missile was aimed. If we think of the center of the target as the origin of a coordinate system, we can let X denote the north–south distance between the landing point and the target center and let Y denote the corresponding east–west distance. (Assume that north and east define positive directions.) The distance between the landing point and the target center is then $U = \sqrt{X^2 + Y^2}$. If X and Y are independent, standard normal random variables, find the probability density function for U.

2. (**Rice 3.66**) Each component of the following system (see below) has an independent exponentially distributed lifetime with parameter λ. Find the cdf and the pdf of the system's lifetime.



3. (**Rice 4.61a**) Let X and Y be continuous random variables with joint pdf given by

$$f(x, y) = 2, \quad 0 \le x \le y \le 1.$$

Identify the marginal distributions of X and Y and then compute $\rho_{X,Y}$, i.e. the correlation coefficient of X and Y. Does the sign of the correlation make sense intuitively?

MH 2500 Probability and introduction to statistics

Tutorial 10 Week 11. Sections 3.6.1, 3.7, 4.3, 4.5. Solutions

1. Let X and Y be i.i.d. $N(0,1)$-distributed random variables. We want to determine the distribution of
$$U = \sqrt{X^2 + Y^2}.$$
Since the square of a standard normal random variable is $\chi^2(1)$, it follows that X^2 and Y^2 are i.i.d. $\chi^2(1)$-distributed random variables. During lecture we proved that the sum of n i.i.d. $\chi^2(1)$-distributed random variables is $\chi^2(n)$, and so it follows that $W = X^2 + Y^2 \sim \chi^2(2)$. We are now interested in the probability distribution of $U = \sqrt{W}$ and since this is not a linear transformation we use the cdf-method. We first notice that the pdf of W is given by
$$f_w(w) = \frac{1}{2}e^{-w/2}, \quad w > 0$$
It now follows that
$$F_U(u) = \Pr(U \le u) = \Pr\left(\sqrt{W} \le u\right) = \Pr\left(W \le u^2\right) = F_w\left(u^2\right), \ u > 0$$
and so
$$\begin{aligned} f_U(u) &= 2u \cdot f_w\left(u^2\right) = 2u \cdot \frac{1}{2}e^{-u^2/2} = \\ &= ue^{-u^2/2}, \quad u > 0 \end{aligned}$$
and we conclude that U is a Rayleigh distributed random variable.

2. This is a system where the components are connected both in series and in parallel Let $X_1, X_2, \ldots, X_6$ represent the lifetimes of each of the individual components, which means that these sex random variables are i.i.d. $Exp(\lambda)$. If we now let V_1, V_2, and V_3 represent the lifetimes of the three parts that are connected in series, it follows that these three random variables are i.i.d. where, for example, $V_1 = \min(X_1, X_2)$. According to Example A in Section 3.7, it follows that V_1, V_2, and V_3 are i.i.d. $Exp(2\lambda)$. If we finally let U represent the lifetime of the entire system, we have that $U = \max(V_1, V_2, V_3) = V_{(3)}$ and so
$$F_U(u) = (F_V(u))^3 = \left(1 - e^{-2\lambda u}\right)^3, \quad u > 0$$
which means that
$$f_U(u) = 6\lambda e^{-2\lambda u}\left(1 - e^{-2\lambda u}\right)^2, \quad u > 0$$

3. Let X and Y be continuous random variables with joint pdf given by

$$f(x,y) = 2, \quad 0 \le x \le y \le 1.$$

Identify the marginal distributions of X and Y and then compute $\rho_{X,Y}$, i.e. the correlation coefficient of X and Y. Does the sign of the correlation make sense intuitively? The marginal pdf of X and Y are given by

$$f_X(x) = \int_x^1 2dy = 2(1-x), \quad 0 \le x \le 1,$$

and

$$f_Y(y) = \int_0^y 2dx = 2y, \quad 0 \le y \le 1.$$

We identify these as pdfs of the beta distribution, i.e. $X \sim Beta(1,2)$ and $Y \sim Beta(2,1)$. The information that we have about the beta distribution tells us that

$$\begin{aligned} E(X) &= 1/3, \quad Var(X) = 1/18, \\ E(Y) &= 2/3, \quad Var(Y) = 1/18. \end{aligned}$$

With the aid of the joint pdf of X and Y we find that

$$E(XY) = \int_0^1 \int_0^y 2xydxdy = \int_0^1 y\left[x^2\right]_0^y dy = \int_0^1 y^3 dy = \frac{1}{4}\int_0^1 4y^3 dy = \frac{1}{4},$$

since the integrand in the last expression is the pdf of $Beta(4,1)$. So, finally,

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{\frac{1}{4} - \frac{1}{3}\cdot\frac{2}{3}}{\sqrt{\frac{1}{18}}\cdot\sqrt{\frac{1}{18}}} = \frac{1}{2}.$$

The fact that the correlation is positive is perfectly logical since, for example, a large value of X must of course result in a large value of Y.

MH 2500 Probability and introduction to statistics

Tutorial 11 Week 12. Section 4.4.1. Problems

1. (**Rice 4.61bc**) Let X and Y be continuous random variables with joint pdf given by
$$f(x, y) = 2, \quad 0 \leq x \leq y \leq 1.$$
   (a) Find $E(X \mid Y = y)$ and $E(Y \mid X = x)$. Do these results make sense intuitively?
   (b) Find the probability density functions of the random variables $E(X \mid Y)$ and $E(Y \mid X)$.

2. (**Rice 4.75**) Let T be an exponential random variable, and conditional on T, let U be uniform on $[0, T]$. Find the unconditional mean and variance of U.

MH 2500 Probability and introduction to statistics

Tutorial 11 Week 12. Section 4.4.1. Solutions

1. Let X and Y be continuous random variables with joint pdf given by

$$f(x, y) = 2, \quad 0 < x < y < 1.$$

Remark. From the previous tutorial (**Rice 4.61a**) we have that $X \sim Beta(1, 2)$ and $Y \sim Beta(2, 1)$.

(a) The conditional pdf of X given $Y = y$ is

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)} = \frac{2}{2y} = \frac{1}{y}, \quad 0 < x < y < 1$$

which means that $X \mid Y = y \sim U(0, y)$ and therefore

$$E(X \mid Y = y) = \frac{y}{2}$$

Furthermore, the conditional pdf of Y given $X = x$ is

$$f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)} = \frac{2}{2(1-x)} = \frac{1}{1-x}, \quad 0 < x < y < 1$$

which means that $Y \mid X = x \sim U(x, 1)$ and so

$$E(Y \mid X = x) = \frac{1+x}{2}$$

Both these conditional distributions/expectations are very logical since after we are given the value of either X or Y, we don't have any other preconceptions about the value of the other variable than that Y must be larger than the given value of X and that X must be smaller than the given value of Y.

(b) Using the previous results, we have that $X \mid Y \sim U(0, Y)$ which means that

$$U = h(Y) = E(X \mid Y) = \frac{Y}{2}$$

Therefore

$$F_U(u) = \Pr(U \le u) = \Pr\left(\frac{Y}{2} \le u\right) = \Pr(Y \le 2u) = F_Y(2u)$$

and so

$$f_U(u) = f_Y(2u) \cdot 2 = 8u, \quad 0 \le u \le \frac{1}{2}$$

Furthermore, $Y \mid X \sim U(1, X)$, i.e.

$$V = h(X) = E(Y \mid X) = \frac{1+X}{2}$$

Therefore

$$F_V(v) = \Pr(V \le v) = \Pr\left(\frac{1+X}{2} \le v\right) = \Pr(X \le 2v-1) = F_X(2v-1)$$

and so

$$f_V(v) = f_X(2v-1) \cdot 2 = 4(1-(2v-1)) = 8(1-v), \quad \frac{1}{2} \le v \le 1$$

2. According to the assumptions, $T \sim Exp(\lambda)$ and $U \mid T \sim U(0, T)$. The formulas for the law of total expectation and the law of total variance yields

$$E(U) = E(E(U \mid T)) = E\left(\frac{T}{2}\right) = \frac{1}{2}E(T) = \frac{1}{2\lambda}$$

and

$$\begin{aligned} Var(U) &= E(Var(U \mid T)) + Var(E(U \mid T)) = E\left(\frac{T^2}{12}\right) + Var\left(\frac{T}{2}\right) = \\ &= \frac{1}{12}\left(Var(T) + (E(T))^2\right) + \frac{1}{4}Var(T) = \\ &= \frac{1}{12}\left(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\right) + \frac{1}{4\lambda^2} = \frac{5}{12\lambda^2} \end{aligned}$$

MH 2500 Probability and introduction to statistics

Tutorial 12 Week 13. Chapter 5. Problems

1. (**Rice 5.1**) Let $X_1, X_2, \ldots$ be a sequence of independent random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma_i^2$, for $i = 1, 2, \ldots$. Furthermore, let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ for $n = 1, 2, \ldots$. Show that if
$$\frac{\sum_{i=1}^{n} \sigma_i^2}{n^2} \to 0 \quad \text{as} \quad n \to \infty,$$
then $\overline{X}_n \xrightarrow{p} \mu$.
Remark. This is a generalization of LLN. A further generalization can be found in Problem 5.2. Try to solve that one too.

2. (**Rice 5.5**) Using moment generating functions, show that as $n \to \infty$, $p \to 0$, and $np \to \lambda$, the Binomial distribution with parameters n and p tends to the Poisson distribution. *Hint.* If $a_n \to a$ as $n \to \infty$ then $(1 + a_n/n)^n \to e^a$ as $n \to \infty$.

3. (**Rice 5.16**) Suppose that $X_1, X_2, \ldots, X_{20}$ are i.i.d. random variables with pdf given by
$$f(x) = 2x, \quad 0 \le x \le 1.$$
Let $S = X_1 + X_2 + \cdots + X_{20}$. Use the Central Limit Theorem to approximate $\Pr(S \le 10)$.

4. A machine that produces wine corks produces corks where the diameter that can be considered to be normally distributed with a mean of 3cm and a standard deviation of 0.1cm. A cork is acceptable if its diameter is between 2.9cm and 3.1 cm. Suppose we have 60 (not yet assessed) corks in stock and that we need 40 acceptable corks for a batch of wine bottles. Use the Central Limit Theorem to approximate the probability that we can cork the bottles without having to let the machine produce new corks.

MH 2500 Probability and introduction to statistics

Tutorial 12 Week 13. Chapter 5. Solutions

1. (**Rice 5.1**) Let $X_1, X_2, \ldots$ be a sequence of independent random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma_i^2$, for $i = 1, 2, \ldots$. Furthermore, let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ for $n = 1, 2, \ldots$. Show that if

$$\frac{\sum_{i=1}^n \sigma_i^2}{n^2} \to 0 \quad \text{as} \quad n \to \infty,$$

then $\overline{X}_n \xrightarrow{p} \mu$. Since

$$\begin{aligned} E\left(\overline{X}_n\right) &= \mu \\ Var\left(\overline{X}_n\right) &= \frac{\sum_{i=1}^n \sigma_i^2}{n^2} \end{aligned}$$

it follows from Chebyshev's inequality that for any $\varepsilon > 0$,

$$\Pr\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) \le \frac{\sum_{i=1}^n \sigma_i^2}{n^2\varepsilon^2} \to 0 \quad \text{as} \quad n \to \infty$$

and we have proved that $\overline{X}_n \xrightarrow{p} \mu$.

2. (**Rice 5.5**) Using moment generating functions, show that as $n \to \infty$, $p \to 0$, and $np \to \lambda$, the Binomial distribution with parameters n and p tends to the Poisson distribution. Let $X_n \sim Bi(n, p_n)$ for $n = 1, 2, 3, \ldots$ where we have that $\lambda_n = np_n \to \lambda$ as $n \to \infty$. Then

$$M_{X_n}(t) = \left[p_n e^t + (1 - p_n)\right]^n = \left[1 + p_n\left(e^t - 1\right)\right]^n = \left(1 + \frac{\lambda_n\left(e^t - 1\right)}{n}\right)^n$$

Since $\lambda_n \to \lambda$ as $n \to \infty$ it follows by the hint that

$$\lim_{n\to\infty} M_{X_n}(t) = \lim_{n\to\infty}\left(1 + \frac{\lambda_n\left(e^t - 1\right)}{n}\right)^n = e^{\lambda\left(e^t - 1\right)}$$

which we recognize as the mgf of $Po(\lambda)$.

3. (**Rice 5.16**) Suppose that $X_1, X_2, \ldots, X_{20}$ are i.i.d. random variables with pdf given by

$$f(x) = 2x, \quad 0 \le x \le 1.$$

Let $S = X_1 + X_2 + \cdots + X_{20}$. Use the Central Limit Theorem to approximate $\Pr(S \le 10)$. Since S is a sum of a relatively large amount of not too asymmetrically distributed i.i.d. random variables, the Central Limit Theorem can be applied. First we need the mean and the variance of this distribution. The easiest way to do this is to make the observation that $X \sim Beta(2,1)$. However, if we do not see this we have to do it the hard way (which in this case isn't so hard). It follows first that

$$E(X) = \int_0^1 2x^2 dx = \frac{2}{3}\left[x^3\right]_0^1 = \frac{2}{3}.$$

Since

$$E\left(X^2\right) = \int_0^1 2x^3 dx = \frac{1}{2}\left[x^4\right]_0^1 = \frac{1}{2}$$

it then follows by the calculation formula for the variance that

$$Var(X) = E\left(X^2\right) - (E(X))^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

Since S is a sum of 20 i.i.d. such random variables we get that

$$\begin{aligned} E(S) &= 20 \cdot \frac{2}{3} = \frac{40}{3} \\ Var(S) &= 20 \cdot \frac{1}{18} = \frac{10}{9} \end{aligned}$$

and so $S \overset{Approx}{\sim} N\left(\frac{40}{3}, \frac{10}{9}\right)$. Using the standardization technique for the normal distribution it follows that

$$\Pr(S \le 10) \approx \Pr\left(Z \le \frac{10 - \frac{40}{3}}{\frac{\sqrt{10}}{3}} = -\sqrt{10} \approx -3.16\right) = 0.0008$$

4. A machine that produces wine corks produces corks where the diameter can be considered to be normally distributed with a mean of 3cm and a standard deviation of 0.1cm. A cork is acceptable if its diameter is between 2.9cm and 3.1cm. Suppose we have 60 (not yet assessed) corks in stock and that we need 40 acceptable corks for a batch of wine bottles. Use the Central Limit Theorem to approximate the probability that we can cork the bottles without having to let the machine produce new corks. Let X represent the diameter of a randomly chosen cork. By the assumptions, $X \sim N(3, 0.01)$. The first we need to do is to determine the probability that a randomly chosen cork is acceptable. If we denote this probability p, then

$$\begin{aligned} p &= \Pr(2.9 \leq X \leq 3.1) = \Pr\left(\frac{2.9-3}{0.1} \leq Z \leq \frac{3.1-3}{0.1}\right) = \Pr(-1 \leq Z \leq 1) = \\ &= \Pr(Z \leq 1) - \Pr(Z \leq -1) = \Pr(Z \leq 1) - \Pr(Z \geq 1) = \\ &= \Pr(Z \leq 1) - [1 - \Pr(Z \leq 1)] = 2 \cdot \Pr(Z \leq 1) - 1 = 2 \cdot 0.8413 - 1 = 0.6826 \end{aligned}$$

Now let Y represent the number of *acceptable* corks in stock. If we find it reasonable to assume that the diameters of different corks are indpendent of each other, then it is clear that $Y \sim Bin(60; 0.6826)$. We are interested in finding $\Pr(Y \geq 40)$. Since both $np = 60 \cdot 0.6826 = 41.0 > 5$ and $n(1-p) = 60 \cdot (1 - 0.6826) = 19.0 > 5$, the rule of thumb given in the textbook tells us that the Central Limit Theorem can be applied in this situation. Because of the fact that

$$\begin{aligned} E(Y) &= np = 60 \cdot 0.6826 = 41.0 \\ Var(Y) &= np(1-p) = 60 \cdot 0.6826 \cdot (1 - 0.6826) = 13.0 \end{aligned}$$

we conclude that $Y \overset{Approx}{\sim} N(41.0, 13.0)$. Therefore,

$$\Pr(Y \geq 40) \approx \Pr\left(Z \geq \frac{40 - 41.0}{\sqrt{13.0}} = -0.28\right) = \Pr(Z \leq 0.28) = 0.61$$

Remark. The exact binomial probability is given by

$$\Pr(Y \geq 40) = \sum_{y=40}^{60} \binom{60}{y} \cdot 0.6826^y \cdot (1 - 0.6826)^{60-y} = 0.662$$

and we see that the approximation is rather poor. Due to the fact that the binomial distribution is discrete, we would get a better approximation by using *continuity correction* (or *half correction*). The approximation would then be

$$\Pr(Y \geq 40) \approx \Pr\left(Z \geq \frac{39.5 - 41.0}{\sqrt{13.0}} = -0.42\right) = \Pr(Z \leq 0.42) = 0.662$$

and all of a sudden, the approximation is close to perfect.