PROBABILITY THEORY
FALL 2019

Time and place: TF2 9:50–11:10am, Room CA-A5, College Avenue Campus
Instructor: Claire Burrin, claire.burrin@rutgers.edu, Hill 232, Busch Campus
Office hours: Every other Wednesday, 1:30-3pm, Hill 232. The precise dates are: Sep 11, Sep 25, Oct 16, Oct 30, Nov 6, Nov 20, Dec 4
Prerequisites: Math 251, esp. working knowledge of multiple integrals and partial derivatives.
Textbook: Sheldon Ross, A first course in probability, 8th ed.
Final: Monday, Dec 16, 12-3pm, Room CA-A5

This is a first course in the mathematical theory of probability, where we will study random events (e.g. will it rain tomorrow), discrete and continuous random variables (e.g. the number of rainy days in September vs. the total quantity of rain in September), and learn to compute their probability, expected value, standard deviation, density function, etc. We will discuss the two fundamental results of probability theory, the Law of Large Numbers (LLN) that says, roughly, that the average value of a large number of experiments is approximatively the expected value, and the Central Limit Theorem (CLT) that determines the distribution of any sum of a large number of iid random variables, regardless of their underlying distribution.

There is a general dichotomy according to which a problem behaves either deterministically (e.g. the solution of an equation) or randomly according to some probabilistic law. The hard part is usually to identify the law. One guiding example will be $\omega(n) = \sum_{p \mid n} 1$, the number of distinct prime divisors of $n$ (where $n$ is chosen at random in a large discrete interval $\{1, \ldots, N\}$). For example, $\omega(3) = 1$, $\omega(4) = 1$, $\omega(5) = 1$, $\omega(6) = 2$, $\omega(100) = 2$. Here is a plot of its values (retrieved from the Wolfram website):

The jumps we are observing do not occur at regular intervals, so is there a probabilistic law that governs this distribution? As an application of CLT, we will prove the Erdős–Kac theorem, which identifies the probabilistic behavior of the function $\omega(n)$.

Evaluation: The grade make-up is 20% Homework, 20% each midterm, 40% the final. There will be no quizzes. Anyone absent from an exam without medical proof will receive a zero score. Exceptional circumstances will require a letter from a Dean. Students are expected to behave in accordance with the Code of Academic Integrity [http://academicintegrity.rutgers.edu/academic-integrity-policy](http://academicintegrity.rutgers.edu/academic-integrity-policy). Cases of cheating will be reported.
Tentative syllabus (will be periodically updated)

<table>
<thead>
<tr>
<th>Week</th>
<th>Topics</th>
<th>Textbook reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep 6</td>
<td>Combinatorics</td>
<td>1.1-1.4</td>
</tr>
<tr>
<td>Sep 10, 13</td>
<td>Combinatorics, axioms of probability</td>
<td>1.5, 2.1-2.5</td>
</tr>
<tr>
<td>Sep 17, 20</td>
<td>Conditional probability, independence</td>
<td>3.1-3.3, 3.5</td>
</tr>
<tr>
<td>Sep 24, 27</td>
<td>Discrete random variables: E[X], Var(X),...</td>
<td>3.4, 4.1-4.5</td>
</tr>
<tr>
<td>Oct 1, 4</td>
<td>Bernoulli, binomial, Poisson distributions</td>
<td>4.5-4.10</td>
</tr>
<tr>
<td>Oct 8, 11</td>
<td>Continuous RVs, Midterm 1</td>
<td>Chap 5.1-2</td>
</tr>
<tr>
<td>Oct 15, 18</td>
<td>Common continuous distributions</td>
<td>5.3-6</td>
</tr>
<tr>
<td>Oct 22, 25</td>
<td>Joint distributions</td>
<td>5.6-7, 6.1</td>
</tr>
<tr>
<td>Oct/Nov 29, 1</td>
<td>Re-conditional probability and independence</td>
<td>6.2-7</td>
</tr>
<tr>
<td>Nov 5, 8</td>
<td>Expectation, covariance, and correlation</td>
<td>7.1-2, 7.4</td>
</tr>
<tr>
<td>Nov 12, 15</td>
<td>Midterm 2 this week</td>
<td></td>
</tr>
<tr>
<td>Nov 19, 22</td>
<td>Conditional expectation, moment generating function and CLT</td>
<td>7.5-7,8.3</td>
</tr>
<tr>
<td>Nov 27</td>
<td>Thanksgiving week</td>
<td></td>
</tr>
<tr>
<td>Dec 3, 6</td>
<td>Inequalities and laws of large numbers</td>
<td></td>
</tr>
<tr>
<td>Dec 10</td>
<td>Additional Topics</td>
<td>Chapter 8</td>
</tr>
</tbody>
</table>

Homework assignments: Homework solutions are due on Tuesdays during the class. Late hand-ins will not be accepted. Remember that your solutions are to be read – and graded! – by someone. Therefore, you should make sure to explain your reasoning as clearly as possible, and to write neatly – this is also considered in the grading. Your homework solutions are expected to show some original effort; solutions that are directly copied from the internet will be sanctioned.

Homework 1, due Sep 17

1. How many different 7-place license plates are possible if the first 2 places are for letters and the other 5 for numbers? And if no letter or number can be repeated in a single plate?
2. John, Jim, Jay, and Jack form a band consisting of 4 instruments. If each boy can play all 4 instruments, how many different arrangements are possible? What if John and Jim can play all 4 instruments, but Jay and Jack can each play only piano and drums?
3. A dance class consists of 22 students, of which 10 are women, and 12 are men. If 5 men and 5 women are to be chosen and then paired off, how many results are possible?
4. A person has 8 friends, of whom 5 will be invited to a party. How many choices are there if 2 of the friends are feuding, and will not attend together? How many choices if 2 of the friends will only attend together?
5. Consider the grid of points below. Suppose that, starting at point A, you can go one step up or one step to the right at each move. This procedure is continued until the point labeled B is reached. How many different paths from A to B are possible?
(6) Give a combinatorial proof that
\[
\binom{n + m}{r} = \binom{n}{0} \binom{m}{r} + \binom{n}{1} \binom{m}{r-1} + \cdots + \binom{n}{r} \binom{m}{0}.
\]

(7) Two (indistinguishable) dice are thrown. Let \( E \) be the event that the sum of the dice is odd, let \( F \) be the event that at least one of the dice lands on 1, and let \( G \) be the event that the sum is 5. Describe the events \( E \cap F, E \cup F, E \cap F^c, E \cap F \cap G \).

(8) The following data was given in a study of a group of 1000 subscribers to a certain magazine. In reference to job, marital status, and education, there were 312 professionals, 470 married persons, 525 college graduates, 42 professional college graduates, 147 married college graduates, 86 married professionals, and 25 married professional college graduates. Show that the numbers reported in the study must be incorrect.

\textit{Hint: Inclusion-exclusion.}

\textbf{Homework 2, due Sep 24}

(1) Prove that \( P(E \cap F^c) = P(E) - P(E \cap F) \).

(2) A bowl contains \( n \) white and \( m \) black balls.
   
   (a) If two balls are randomly withdrawn, what is the probability that they are the same color?
   
   (b) If a ball is randomly withdrawn and then replaced before the second one is drawn, what is the probability that the withdrawn balls are the same color?
   
   (c) Show that the probability in part (b) is always larger than the one in part (a).

(3) \( b \) boys and \( g \) girls are lined up in random order; that is, each of the \( (b+g)! \) permutations is assumed to be equally likely. What is the probability that the person in the \( i \)-th position, for \( 1 \leq i \leq (b+g) \), is a girl?

(4) Three cards are randomly selected, without replacement, from a standard deck of 52 cards. Compute the conditional probability that the first card selected is a spade given that the second and third cards are spades.

(5) If 3 married couples are arranged in a row, find the probability that no husband sits next to his wife.

(6) Consider an experiment whose sample space consists of a countably infinite number of points. Show that not all points can be equally likely. Can all points have a positive probability of occurring?

(7) An investor owns shares in a stock whose present value is 25. She has decided that she must sell her stock if it goes either down to 10 or up to 40. If each change of price is either up 1 point with probability 0.55 or down 1 point with probability 0.45, and the successive changes are independent, what is the probability that the investor retires a winner?

\textbf{Homework 3, due Oct 1}
(1) The following method was proposed to estimate the number of people over the age of 50 who reside in a town of known population 100,000: As you walk along the streets, keep a running count of the percentage of people you encounter who are over 50. Do this for a few days; then multiply the percentage you obtain by 100,000 to obtain the estimate. Let $p$ denote the proportion of people in the town who are over 50. Furthermore, let $a_1$ denote the proportion of time that a person under the age of 50 spends in the streets, and let $a_2$ be the corresponding value for those over 50. What quantity does the method suggested estimate? When is the estimate approximately equal to $p$?

(2) Prove the following equalities: for any sets $A_1, \ldots, A_n$,
\[
\left( \bigcup_{k=1}^{n} A_k \right)^c = \bigcap_{k=1}^{n} A_k^c, \quad \left( \bigcap_{k=1}^{n} A_k \right)^c = \bigcup_{k=1}^{n} A_k^c.
\]

(3) A parallel system functions whenever at least one of its components works. Consider a parallel system of $n$ components, and suppose that each component works independently with probability $1/2$. Find the conditional probability that component 1 works given that the system is functioning.

(4) An insurance company has the following policy: $A$ must be paid if event $E$ happens within a year. The company predicts that the probability of $E$ happening within a year is $p$. How much should the company charge the customer so that its expected profit is 10 percent of $A$?

(5) On a multiple-choice exam with 3 possible answers for each of the 5 questions, what is the probability that a student will get 4 or more correct answers just by guessing?

(6) Suppose that, in flight, airplane engines will fail with probability $1 - p$, independently from engine to engine. If an airplane needs a majority of its engines operative to complete a successful flight, for what values of $p$ is a 5-engine plane preferable to a 3-engine plane?

Homework 4, due Oct 8

(1) A coin that, when flipped, comes up heads with probability $p$ is flipped until either heads or tails has occurred twice. Find the expected number of flips.

(2) If $X$ is a binomial random variable with expected value 6 and variance 2.4, find $P(X = 5)$.

(3) On average, 5.2 hurricanes hit a certain region in a year. What is the probability that there will be 3 or fewer hurricanes hitting this year?

(4) Approximately 80,000 marriages took place in New York state last year. What is the probability that, for at least one couple,
   (a) both partners were born on April 30?
   (b) both partners celebrated their birthday on the same day of the year?

(5) A roulette wheel consists of 38 numbers: 1, \ldots, 36, 0, 00. If Bob always bets that the outcome will be in $\{1, \ldots, 12\}$, what is the chance that
   (a) he loses his first 5 bets?
   (b) his first win occurs on his fourth bet?

(6) Let $X$ be a random variable with expected value $\mu$ and variance $\sigma^2$. What is the expected value and variance of $Y = \frac{X - \mu}{\sigma}$?

Homework 5, due Oct 22

(1) The number of minutes of playing time of a certain high school basketball player in a randomly chosen game is a random variable whose probability density function is given in the following figure:
(1) A fire station is to be located along a road of length $A$, $A < q$. If fires occur at points uniformly chosen on $(0, A)$, where should the station be located so as to minimize the expected distance from the fire? That is, choose $a$ so as to minimize $E[|X - a|]$ when $X$ is uniformly distributed over $(0, A)$.

(2) If $X$ is uniformly distributed over $(-1, 1)$, find
   
   (a) $P(|X| > 1/2)$;
   (b) the density function of the random variable $|X|$. 

(3) Find the distribution of $R = A \sin \theta$, where $A$ is a fixed constant and $\theta$ is uniformly distributed on $(-\pi/2, \pi/2)$. Such a random variable $R$ arises in the theory of ballistics. If a projectile is fired from the origin at an angle $\alpha$ from the earth with a speed $v$, then the point $R$ at which it returns to the earth can be expressed as $R = (v^2/g) \sin(2\alpha)$, where $g$ is the gravitational constant, equal to 980 centimeters per second squared.

(4) The annual rainfall in Cleveland, Ohio is approximately a normal random variable with mean 40.2 inches and standard deviation 8.4 inches. What is the probability that
   
   (a) next year’s rainfall will exceed 44 inches?
   (b) the yearly rainfalls in exactly 3 of the next 7 years will exceed 44 inches? Assume that if $A_i$ is the event that the rainfall exceeds 44 inches in year $i$ (from now), then the events $A_i$, $i \geq 1$, are independent.

(5) Evidence concerning the guilt or innocence of a defendant in a criminal investigation can be summarized by the value of an exponential random variable $X$ whose mean $\mu$ depends on
whether the defendant is guilty. If innocent, \( \mu = 1 \); if guilty, \( \mu = 2 \). The deciding judge will rule the defendant guilty if \( X > c \) for some suitably chosen value of \( c \).
(a) If the judge wants to be 95 percent certain that an innocent man will not be convicted, what should be the value of \( c \)?
(b) Using the value of \( c \) found in part (a), what is the probability that a guilty defendant will be convicted?

**Homework 7, due Nov 5**

(1) The joint density of \( X \) and \( Y \) is
\[
f(x,y) = \begin{cases} 
  C(y-x)e^{-y} & -y < x < y, \ 0 < y < \infty \\
  0 & \text{otherwise.}
\end{cases}
\]
(a) Find \( C \).
(b) Find the density function of \( X \).
(c) Find the density function of \( Y \).
(d) Find \( E[X] \).
(e) Find \( E[Y] \).
(2) Let \( X_1, \ldots, X_n \) be independent exponential random variables having a common parameter \( \lambda \). Determine the distribution of \( \min(X_1, \ldots, X_n) \).
(3) Three points \( X_1, X_2, X_3 \) are selected at random on a line \( L \). What is the probability that \( X_2 \) lies between \( X_1 \) and \( X_3 \)?
(4) The joint density function of \( X \) and \( Y \) is
\[
f(x,y) = \begin{cases} 
  x + y & 0 < x < 1, \ 0 < y < 1 \\
  0 & \text{otherwise.}
\end{cases}
\]
(a) Are \( X \) and \( Y \) independent?
(b) Find the density function of \( X \).
(c) Find \( P(X + Y < 1) \).
(5) Suppose that \( 10^6 \) people arrive at a service station at times that are independent random variables, each of which is uniformly distributed over \((0,10^6)\). Let \( N \) denote the number that arrive in the first hour. Find an approximation for \( P(N = i) \).

**Homework 8, due Nov 12**

(1) The joint density of \( X \) and \( Y \) is
\[
f(x,y) = c(x^2 - y^2)e^{-x}
\]
for \( 0 < x < \infty, -x \leq y \leq y \). Find the conditional distribution of \( Y \), given \( X = x \).
(2) Let \( X \) and \( Y \) denote the coordinates of a point uniformly chosen in the circle of radius 1 centered at the origin. That is, their joint density is
\[
f(x,y) = \frac{1}{\pi}
\]
for \( x^2 + y^2 \leq 1 \). Find the joint density function of the polar coordinates \( R = (X^2 + Y^2)^{1/2} \) and \( \Theta = \tan(Y/X) \).
(3) An insurance company supposes that each person has an accident parameter and that the yearly number of accidents of someone whose accident parameter is \( \lambda \) is Poisson distributed with mean \( \lambda \). They also suppose that the parameter value of a newly insured person can be assumed to be the value of a gamma random variable with parameters \( s \) and \( \alpha \). If a newly insured person has \( n \) accidents in her first year, find the conditional density of her
accident parameter. Also, determine the expected number of accidents that she will have in the following year.

(4) If $X$ and $Y$ are independent and identically distributed with mean $\mu$ and variance $\sigma^2$, find $E[(X - Y)^2]$.

(5) A deck of $n$ cards numbered 1 through $n$ is thoroughly shuffled so that all possible $n!$ orderings can be assumed to be equally likely. Suppose you are to make $n$ guesses sequentially, where the $i$-th one is a guess of the card in position $i$. Let $N$ denote the number of correct guesses.

(a) If you are not given any information about your earlier guesses show that, for any strategy, $E[N] = 1$.

(b) Suppose that after each guess you are shown the card that was in the position in question. What do you think is the best strategy? Show that, under this strategy,

$$E[N] = \frac{1}{n} + \frac{1}{n-1} + \cdots + 1 \approx \int_1^n \frac{dx}{x} = \log n.$$ 

(c) Suppose that you are told after each guess whether you are right or wrong. In this case, it can be shown that the strategy which maximizes $E[N]$ is one that keeps on guessing the same card until you are told you are correct and then changes to a new card. For this strategy, show that

$$E[N] = 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} \approx e - 1.$$ 

(6) If a dice is to be rolled until all sides have appeared at least once, find the expected number of times that outcome 1 appears.