MA 519: Homework, Midterms and Practice Problems Solutions

Carlos Salinas

Last compiled: December 26, 2016

Contents

1 Homework Solutions

These are my (corrected) solutions to DasGupta's Math/Stat 519 homework for the fall semester of 2016.

Our main references are [1] and [2].

Throughout this document, unless otherwise specified, Ω denotes the sample space in question. The symbol '$\sim$' is used both to denote the distribution type of a random variable and to denote asymptotic equivalence; i.e., if $\{a_n\}$, $\{b_n\}$ are convergent sequences with limit a and b, respectively, we say $a_n \sim b_n$ if $\frac{a_n}{b_n} \to 1$.

Although I am not fond of the following notation, for the sake of remaining consistent with DasGupta's book, we shall adopt the following notation regarding PMFs and PDFs throughout this document (adapting qual problems if necessary):

Name	Symbol	mass/density function	mean	variance	MGF
Bernoulli	$\mathrm{Ber}(p)$	$1-p$ for $k=0$, p for $k=1$	p	$p(1-p)$	$(1-p)+p\mathrm{e}^t$
binomial	$\mathrm{Bin}(n,p)$	$\binom{n}{k}p^k(1-p)^{n-k}$ for $k=0,\dots,n$	np	$np(1-p)$	$(p\mathrm{e}^t+1-p)^n$
Poisson	$\mathrm{Poi}(p)$	$\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ	$\mathrm{e}^{\lambda(\mathrm{e}^t-1)}$
geometric	$\mathrm{Geo}(p)$	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)\mathrm{e}^t}$
negative binomial	$NB(r,p)$	$\binom{k+r-1}{k}(1-p)^r p^k$	$\frac{pr}{1-p}$	$\frac{1-p}{p^2}$	$\frac{pr}{(1-p)^2}$
hypergeometric	$\mathrm{Hypergeo}(n,N,D)$	$\frac{\binom{D}{k}\binom{N-D}{n-x}}{\binom{N}{n}}$	$\frac{nD}{N}$	$\frac{nD}{N}\left(1-\frac{D}{N}\right)\frac{N-n}{N-1}$	
uniform	$U[a,b]$	$\frac{1}{b-a}$ for $a \le x \le b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{\mathrm{e}^{bt}-\mathrm{e}^{at}}{(b-a)t}$
exponential	$\mathrm{Exp}(\lambda)$	$\frac{1}{\lambda}\mathrm{e}^{-x/\lambda}$ for $x \ge 0$	λ	λ^2	$(1-\lambda t)^{-1}$
gamma	$G(\alpha,\lambda)$	$\frac{\mathrm{e}^{-x/\lambda}x^{\alpha-1}}{\lambda^\alpha\Gamma(\alpha)}$ for $x \ge 0$	$\alpha\lambda$	$\alpha\lambda^2$	$(1-\lambda t)^{-\alpha}$
chi-squared	χ^2_n	$\frac{\mathrm{e}^{-x/2}x^{m/2-1}}{2^{m/2}\Gamma(\frac{m}{2})}$ for $x \ge 0$	m	$2m$	$(1-2t)^{-m/2}$
beta	$B(\alpha,\beta)$	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ for $0 \le x \le 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	
normal	$N(\mu,\sigma)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(x-\mu)^2/2\sigma^2}$	μ	σ^2	$\mathrm{e}^{t\mu+t^2\sigma^2/2}$
Cauchy	$\mathrm{Cauchy}(\mu,\sigma)$	$\left(\sigma\pi\left(1+\frac{(x-\mu)^2}{\sigma^2}\right)\right)^{-1}$	DNE	DNE	DNE

For numerical problems, we consistently stick to the rule of writing four figures after the decimal point.

1.1 Homework 1

PROBLEM 1.1.1 (Handout 1, # 5). A closet contains five pairs of shoes. If four shoes are selected at random, what is the probability that there is at least one complete pair among the four?

SOLUTION. Let $A \subset \Omega$ denote the event "there is at least one complete pair of shoes among the four randomly selected shoes."

First, we compute the size of the sample space Ω. There are $\binom{10}{4} = 210$ ways to select four shoes from five pairs of shoes. Therefore, $|\Omega| = 210$ so each sample point $x \in \Omega$ has an associated probability of $P(x) = \frac{1}{210}$.

Now, let us compute $P(A)$. Oftentimes it is easier to compute the probability $P(\Omega \setminus A)$ and use the identity

$$P(A) + P(\Omega \setminus A) = 1;$$

this is one of those times. Let us now find the number of sample points in $\Omega \setminus A$; i.e., the event "there is no complete pair of shoes among the four randomly selected shoes." In order that we do not choose a complete pair of shoes we must choose from four different pairs, this can be done in $\binom{5}{4} = 5$ ways, and for each pair we have the possibility of choosing on of two shoes belonging to that pair (either a left or a right shoe). Therefore, $|\Omega \setminus A| = 5 \cdot 2^4 = 80$ and so

$$\begin{aligned} P(A) &= 1 - P(\Omega \setminus A) \\ &= 1 - \frac{80}{210} \\ &\approx 0.619. \end{aligned} \qquad \blacksquare$$

PROBLEM 1.1.2 (Handout 1, # 7). A gene consists of 10 subunits, each of which is normal or mutant. For a particular cell, there are 3 mutant and 7 normal subunits. Before the cell divides into two daughter cells, the gene duplicates. The corresponding gene of cell 1 consists of 10 subunits chosen from the 6 mutant and 14 normal units. Cell 2 gets the rest. What is the probability that one of the cells consists of all normal subunits.

SOLUTION. Let $A \subset \Omega$ denote the event "at least one daughter cell contains all normal subunits."

When the cell duplicates, there will be a total of 20 subunits (14 normal and 6 mutant ones). There are $\binom{20}{10} = 184756$ ways to distribute these subunits to a given daughter cell. Therefore, $|\Omega| = 184756$.

Now, let us count the number of sample points in A. Suppose, with out loss of generality, that cell 1 receives all of the normal subunits; there are $\binom{14}{10} = 1001$ ways to do this. Since we may as well have chosen cell 2 to give the normal units to, the number of sample points in A is twice the figure above; i.e., $|A| = 2002$.

Therefore,

$$P(A) = \frac{2002}{184756} \approx 0.0108. \qquad \blacksquare$$

PROBLEM 1.1.3 (Handout 1, # 9). From a sample of size n, r elements are sampled at random. Find the probability that none of the N prespecified elements are included in the sample, if sampling is

(a) with replacement;
(b) without replacement.

Compute it for $r = N = 10$, $n = 100$.

SOLUTION. For part (a): The size of the sample space is $|\Omega| = n^r$ so for each $x \in \Omega$, $P(x) = \frac{1}{n^r}$. If n elements are prespecified, there are $n - N$ non-prespecified elements and thus we have $(n - N)^r$ ways to draw r non-prespecified elements. Thus, the probability that none of the N prespecified elements are drawn if we sample r elements randomly with replacement is

$$p_1(n, N, r) = \frac{(n-N)^r}{n^r}. \qquad (1.1.1)$$

For part (b): The argument leading to the probability of this event is similar to that of part (a). The size of the sample space is $|\Omega| = \binom{n}{r}$; these correspond to the possible draws of r elements without replacement. As before, $n - N$ of the elements have not been prespecified and therefore, we have $\binom{n-N}{r}$ ways of drawing r of the non-prespecified elements. Thus, the probability that note of the N prespecified elements are drawn if we sample r elements randomly without replacement is

$$p_2(n, N, r) = \frac{\binom{n-N}{r}}{\binom{n}{r}}. \tag{1.1.2}$$

Finally, we compute p_1 and p_2 for $r = N = 10$, $n = 100$ using equations (1.1.1) and (1.1.2) above:

$$p_1(100, 10, 10) \approx 0.3487,$$
$$p_2(100, 10, 10) \approx 0.3305. \qquad \blacksquare$$

Problem 1.1.4 (Handout 1, # 11). Let E, F, and G be three events. Find expressions for the following events:

(a) only E occurs;
(b) both E and G occur, but not F;
(c) all three occur;
(d) at least one of the events occurs;
(e) at most two of them occur.

Solution. For part (a): The event "from E, F, and G only E occurs" is given by

$$E \cap (\Omega \setminus F) \cap (\Omega \setminus G).$$

For part (b): The event "from E, F, and G both E and G occur, but not F" is given by

$$E \cap G \cap (\Omega \setminus F).$$

For part (c): The event "from E, F, and G all three occur" is given by

$$E \cap F \cap G.$$

For part (d): The event "from E, F, and G at least one occurs" is given by

$$E \cup F \cup G.$$

For part (e): The event "from E, F, and G at most two occur" is given by

$$\big(E \cap F \cap (\Omega \setminus G)\big) \cup \big(E \cap (\Omega \setminus F) \cap G\big) \cup \big((\Omega \setminus E) \cap F \cap G\big). \qquad \blacksquare$$

PROBLEM 1.1.5 (Handout 1, # 12). Which is more likely:

(a) Obtaining at least one six in six rolls of a fair die, or
(b) Obtaining at least one double six in six rolls of a pair of fair dice.

SOLUTION. For part (a): The probability that we do not roll a six in six rolls of a fair die is

$$q_1 = \left(\frac{5}{6}\right)^6 \approx 0.3349.$$

Therefore, the probability of seeing at least one six in six rolls of a fair die is

$$p_1 = 1 - q_1 \approx 0.6651.$$

For part (b): The probability that we do not roll a double six in six rolls of a pair fair die is

$$q_2 = \left(\frac{35}{36}\right)^6 \approx 0.84449.$$

Therefore, the probability of seeing at least one double six in six rolls of a pair of fair die is

$$p_2 = 1 - q_2 \approx 0.1555.$$

Lastly, we see that the $p_1 > p_2$; i.e., the probability of rolling at least one six in six rolls of a fair die is more likely than the probability of obtaining two double sixes is six rolls of a pair of fair dice. $\blacksquare$

PROBLEM 1.1.6 (Handout 1, # 13). There are n people are lined up at random for a photograph. What is the probability that a specified set of r people happen to be next to each other?

SOLUTION. The r prespecified people can stand as a group starting at positions $1, \ldots, n-r+1$. They can be permuted among themselves in $r!$ ways. The remaining $n-r$ people can be permuted among themselves in $(n-r)!$ ways. Thus, we have

$$p = \frac{(n-r+1)r!(n-r)!}{n!} = \frac{(n-r+1)!r!}{n!}. \qquad \blacksquare$$

PROBLEM 1.1.7 (Handout 1, # 16). Consider a particular player, say North, in a Bridge game. Let X be the number of aces in his hand. Find the distribution of X.

Solution. Let X denote the number of aces in North's hand; this is a random variable taking integer values between 0 and 4. We are asked to find the PMF of X; i.e., the values $P(X = x)$ for all $x = 0, \dots, 4$.

From a deck of 52 cards, 13 cards can be selected in $\binom{52}{13}$ ways. From these North can have x number of aces in $\binom{4}{x}\binom{48}{13-x}$ ways. Therefore, the PMF of X is precisely

$$P(X = x) = \frac{\binom{4}{x}\binom{48}{13-x}}{\binom{52}{13}}.$$

The values of P at each $x = 0, \dots, 4$ are

$$\begin{aligned} &P(X = 0) \approx 0.304, && P(X = 1) \approx 0.439, \\ &P(X = 2) \approx 0.213, && P(X = 3) \approx 0.041, \\ &P(X = 4) \approx 0.003. \end{aligned}$$

■

Problem 1.1.8 (Handout 1, # 20). If 100 balls are distributed completely at random into 100 cells, find the expected value of the number of empty cells.
Replace 100 by n and derive the general expression. Now approximate it as n tends to ∞.

Solution. Let X denote the number of empty cells. Define $I_1, \dots, I_n$ indicator variables as

$$I_k := \begin{cases} 1 & \text{if the } k^{\text{th}} \text{ cell is empty,} \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$X = \sum_{k=1}^{n} I_k$$

so the mean of X is

$$\begin{aligned} E(X) &= \sum_{k=1}^{n} E(I_k) \\ &= \sum_{k=1}^{n} P(I_k = 1) \\ &= \sum_{k=1}^{n} \frac{(n-1)^n}{n^n} \\ &= n\left(1 - \frac{1}{n}\right)^n, \end{aligned}$$

which approaches ∞ as $n \to \infty$ since given any positive real number M, we have

$$\begin{aligned} M &< n\left(1 - \frac{1}{n}\right)^{1/n} \\ 1 &> \left(\frac{M}{n}\right)^{1/n} + \frac{1}{n} \end{aligned}$$

for sufficiently large n.

For $n = 100$, we have $E(X) \approx 36.6032$. ■

1.2 Homework 2

Problem 1.2.1 (Handout 2, # 5). Four men throw their watches into the sea, and the sea brings each man one watch back at random. What is the probability that at least one man gets his own watch back?

Solution. Suppose four men throw their watches into the sea. Place these in some order and label them (from left to right) "the k^{th} man", $k = 1, \ldots, 4$. Let A_k denote the event "the k^{th} man gets his own watch back." Then the event A that at least one man gets his own watch back is the union of these events; i.e., $A = A_1 \cup A_2 \cup A_3 \cup A_4$. Thus, by the inclusion-exclusion principle, we have

$$\begin{aligned} P(A) = P(A_1) + P(A_2) + P(A_3) + P(A_4) \\ &- P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_1 \cap A_4) \\ &- P(A_2 \cap A_3) - P(A_2 \cap A_4) - P(A_3 \cap A_4) \\ &+ P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) \\ &+ P(A_1 \cap A_3 \cap A_4) + P(A_2 \cap A_3 \cap A_4) \\ &- P(A_1 \cap A_2 \cap A_3 \cap A_4). \end{aligned}$$

Since $P(A_k) = P(A_j)$, $P(A_k \cap A_j) = P(A_k \cap A_l)$, etc., for k, j, and l distinct, the equation above reduces to

$$P(A) = 4P(A_1) - 6P(A_1 \cap A_2) + 4P(A_1 \cap A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3 \cap A_4);$$

the choice of A_1, $A_1 \cap A_2$, $A_1 \cap A_2 \cap A_3$, etc., above was arbitrary.

All we need to do now is fill in the blanks. The probability that the 1^{st} man gets back his own wallet is $\frac{1}{4}$ since only one wallet is his own out of the 4. Now, the probability that the 1^{st} and the 2^{nd} man get their own wallet back is $\frac{1}{4} \cdot \frac{1}{3} = \frac{1}{2}$. Proceeding in this fashion, we have

$$P(A) = 4 \cdot \tfrac{1}{4} - 6 \cdot \tfrac{1}{12} + 4 \cdot \tfrac{1}{24} - \tfrac{1}{24} = 0.625. \qquad \blacksquare$$

Problem 1.2.2 (Handout 2, # 7). Calculate the probability that in Bridge, the hand of at least one player is void in a particular suit.

Solution. Order the players and label them "player k", where $k = 1, \ldots, 4$. Let A_k denote the event "player k is void in a prespecified suit." Then the event A "at least one player is void in a prespecified suit" is the union of these events. By the inclusion-exclusion principle (and because of the symmetry of these evenets) we can decompose the computation of $P(A)$ to

$$P(A) = 4P(A_1) - 6P(A_1 \cap A_2) + 4P(A_1 \cap A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3 \cap A_4).$$

Let us now fill in the blanks in the equation above. The probability that player 1 is void in a particular suit, say red clubs ♣, is

$$P(A_1) = \frac{\binom{52-13}{13}}{\binom{52}{13}}.$$

Similarly, the probability that player 1 and player 2 is void in ♣ is

$$P(A_1) = \frac{\binom{52-13}{13}}{\binom{52}{13}} \cdot \frac{\binom{52-13-13}{13}}{\binom{52-13}{13}} = \frac{\binom{52-13-13}{13}}{\binom{52}{13}};$$

and so on.

Thus,

$$P(A) = 4 \cdot \frac{\binom{52-13}{13}}{\binom{52}{13}} - 6 \cdot \frac{\binom{52-13-13}{13}}{\binom{52}{13}} + 4 \cdot \frac{\binom{52-13-13-13}{13}}{\binom{52}{13}} - 1 \cdot 0 \approx 0.05107. \qquad \blacksquare$$

Problem 1.2.3 (Handout 2, # 12). If n balls are placed at random into n cells, find the probability that exactly one cell remains empty.

Solution. There are n ways to choose a cell to be left empty. Of the remaining $n-1$ cells, one must contain 2 balls. That cell can be chosen in $n-1$ ways and the two balls to be placed in it can be chosen in $\binom{2}{n}$ ways. Of the remaining cells, each must contain 1 ball and this pairing can be done in $(n-2)!$ ways. Thus, the probability of the event A that exactly one cell remains empty is

$$P(A) = \frac{n(n-1)\binom{2}{n}(n-2)!}{n^n} = \frac{\binom{2}{n}n!}{n^n}. \qquad \blacksquare$$

Problem 1.2.4 (Handout 2, # 13 – *Spread of rumors*). In a town of $n+1$ inhabitants, a person tells a rumor to a second person, who in turn repeats it to a third person, etc. At each step the recipient of the rumor is chosen at random from the n people available. Find the probability that the rumor told r times without:

(a) returning to the originator,
(b) being repeated to any person.

Do the same problem when at each step the rumor told by one person to a gathering of N randomly chosen people. (The first question is the special case $N = 1$).

Solution. For part (a): The originator can tell any one of the other n inhabitants the rumor. In turn, the non-originator can tell the other $n-1$ inhabitants (not including the originator). Thus, the probability of the event A that the rumor told r times does not return to the originator is

$$P(A) = \frac{n(n-1)^{r-1}}{n^r} = \left(\frac{n-1}{n}\right)^{r-1}.$$

For part (b): The probability of B that the rumor told r times is not repeated to any person is

$$P(B) = \frac{n(n-1)\dots(n-r+1)}{n^r} = \frac{(n)_r}{n^r};$$

the originator is allowed to tell the rumor to any of the other n people while at the k^{th} step, the person telling the rumor is only allowed to tell the rumor to $n-k$ people.

Let us calculate part (a) and (b) now for a rumor told to a gathering of N people.

For part (a): The originator can tell the rumor to any group of N people which can be done in $\binom{n}{N}$ ways while at the k^{th} step, the person can tell the rumor to any group of N people not including the originator; this can be done in $\binom{n-1}{N}$ ways. Therefore, the probability of A_N that the rumor does not return to the originator if it is told to a gathering of N people is

$$P(A_N) = \frac{\binom{n}{N}\binom{n-1}{N}^{r-1}}{\binom{n}{N}^r} = \left(\frac{\binom{n-1}{N}}{\binom{n}{N}}\right)^{r-1}$$

For part (b): The originator can tell the rumor to any group of N people; this can be done in $\binom{n}{N}$ ways. At the k^{th} step, the person can tell the rumor to any group of N people not including the previous $k-1$ people; this can be done in $\binom{n-k+1}{N}$ ways. Therefore, the probability of B_N that the rumor is not repeated to anybody who has already heard it is if it is told to a gathering of N people is

$$P(B_N) = \frac{\binom{n}{N}\cdots\binom{n-r+1}{N}}{\binom{n}{N}^r}.$$

(This equation breaks down for $n - r + 1 < N$.) ∎

Problem 1.2.5 (Handout 2, # 14). What is the probability that

(a) the birthdays of twelve people will fall in twelve different calendar months (assume equal probabilities for the twelve months),
(b) the birthdays of six people will fall in exactly two calendar months?

Solution. For part (a): Let A denote the event that the birthdays of twelve people will fall in twelve different calendar months. There are 12! ways to assign a person to a calendar date without duplication. Therefore,

$$P(A) = \frac{12!}{12^{12}} = \frac{479001600}{8916100448256} \approx 5.3723 \times 10^{-5}$$

For part (b): Let B denote the event that the birthdays of six people will fall in exactly two calendar months. First, there are $\binom{12}{2} = 66$ ways to choose the two calendar months on which the six people's birthdays will fall. For each of the two months, we can have one person having his birthday on the first and eleven on the other; two persons having their birthday on the first and ten on the other; etc. and vice-versa. Therefore,

$$P(B) = \frac{\binom{12}{2}\cdot\binom{2}{1}\cdot\left[\binom{6}{1}+\cdots+\binom{6}{5}\right]}{12^6} \approx 0.002741.$$ ∎

Problem 1.2.6 (Handout 2, # 15). A car is parked among N cars in a row, not at either end. On his return the owner finds exactly r of the N places still occupied. What is the probability that both neighboring places are empty?

Solution. Let A denote the event that upon returning the car owner finds the parking spots neighboring his car empty. Let us count all of the possible arrangements that leave the two spots adjacent to the car owner's car empty. There were originally $N-1$ cars (excluding the car owner's car) so there are $\binom{N-1}{r-1}$ possible arrangements for the remaining cars (excluding the car owner's car); i.e., $|\Omega| = \binom{N-1}{r-1}$. Now, since we want to count those arrangements that leave the two spots adjacent to the car owner's car empty, two of the choices above are forced on us. Therefore, the number of arrangements which leave the two spots adjacent to the car owner's car empty is $\binom{N-1-2}{r-1}$. Thus,

$$P(A) = \frac{\binom{N-3}{r-1}}{\binom{N-1}{r-1}}. \qquad \blacksquare$$

Problem 1.2.7 (Handout 2, # 16). Find the probability that in a random arrangement of 52 bridge card no two aces are adjacent.

Solution. Let A denote the event "there are no two aces adjacent in a random arrangement of 52 bridge card." We can arrange 52 cards in 52! ways so $|\Omega| = 52!$. Now, we can arrange four aces in any of 4! ways and, so that we do not place any ace adjacent to one another, place them in $\binom{48+1}{4} = \binom{49}{4}$ slots in the original 52 set of cards (including the end places). The rest of the 48 cards can be arranged in 48! ways. Thus,

$$P(A) = \frac{\binom{49}{4}4!48!}{52!} \approx 0.7826. \qquad \blacksquare$$

Problem 1.2.8 (Handout 2, # 17). Suppose $P(A) = \frac{3}{4}$, and $P(B) = \frac{1}{3}$.
Prove that $P(A \cap B) \geq \frac{1}{12}$. Can it be equal to $\frac{1}{12}$?

Solution. By the inclusion-exclusion principle, we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

so

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

But, since $A \cup B \subset \Omega$ and $P(\Omega) = 1$, $P(A \cup B) \leq P(\Omega) = 1$ so $-P(A \cup B) \geq -1$. Thus,

$$\begin{aligned}
P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\
&\geq P(A) + P(B) - 1 \\
&= \frac{3}{4} + \frac{1}{3} - 1 \\
&= \frac{9}{12} + \frac{4}{12} - \frac{12}{12} \\
&= \frac{1}{12}.
\end{aligned}$$

Lastly, we show that $P(A \cap B)$ can in fact be equal to $\frac{1}{12}$. Consider the interval $\Omega := [0, 1]$ equipped with the probability measure $P(I) = b - a$ where I is any interval with endpoints $\{a, b\}$ contained in Ω. Set $A := (0, \frac{3}{4})$ and $B := (1 - \frac{1}{3}, 1) = (\frac{2}{3}, 1)$. Then the intersection $A \cap B = (\frac{2}{3}, \frac{3}{4})$ has probability

$$P(A \cap B) = \frac{3}{4} - \frac{2}{3} = \frac{1}{12}. \qquad \blacksquare$$

Problem 1.2.9 (Handout 2, # 18). Suppose you have infinitely many events $A_1, A_2, \ldots$, and each one is sure to occur, i.e., $P(A_k) = 1$ for any i.
Prove that $P\left(\bigcap_{k=1}^{n} A_k\right) = 1$.

Solution. Consider the sequence of probabilities $\{P_n\}$ where $P_n = P\left(\bigcap_{i=1}^{n} A_i\right)$. Note that $\bigcap_{i=1}^{n} A_i \downarrow \bigcap_{i=1}^{\infty} A_i$. First we show, by induction, that $P_n = 1$.

The case $n = 1$ is trivial. Now, assume the result holds for $n - 1$ and consider $P_n = P\left(\bigcap_{i=1}^{n} A_i\right)$. Writing $A' = \bigcap_{i=1}^{n-1} A_i$, we have

$$P_n = P(A' \cap A_n).$$

By the inclusion-exclusion principle,

$$\begin{aligned} P_n &= P(A') + P(A_n) - P(A' \cup A_n) \\ &= 1 + 1 - P(A' \cup A_n) \end{aligned}$$

and since $P(A' \cup A_n) \geq P(A') = 1$ by the monotonicity of the probability measure, $P(A' \cup A_n) = 1$ since $P(A' \cup A_n) \leq P(\Omega) = 1$, thus

$$\begin{aligned} &= 1 + 1 - 1 \\ &= 1. \end{aligned}$$

It follows that $\{P_n\}$ is the constant sequence $\{1\}$ therefore, its limit is 1. $\blacksquare$

Problem 1.2.10 (Handout 2, # 19). There are n blue, n green, n red, and n white balls in an urn. Four balls are drawn from the urn with replacement. Find the probability that there are balls of at least three different colors among the four drawn.

Solution. Let A denote the event "there are balls of at least three different colors among the four drawn." There are 4^4 ways to draw four balls; therefore, $|\Omega| = 4^4$. There are $4!$ wasy to draw a ball of each color and $4 \cdot 3\binom{4}{2}2!$ ways to draw four balls missing exactly one color (pic a color to be missed; pick a color to be drawn twice; pick two draws for that color to be drawn on; and rearrange the other two colors into the other two draws). Thus,

$$P(A) = \frac{4! + 4 \cdot 3\binom{4}{2}2!}{4^4} \approx 0.65625. \qquad \blacksquare$$

1.3 Homework 3

PROBLEM 1.3.1 (Handout 3, # 3). n sticks are broken into one short and one long part. The $2n$ parts are then randomly paired up to form n new sticks. Find the probability that

(a) the parts are joined in their original order, i.e., the new sticks are the same as the old sticks;
(b) each long part is paired up with a short part.

SOLUTION. For part (a): We use the hierarchical probability formula to find the desired probability. Let A_k, where $k = 1, \ldots, n$, denote the event "k^{th} time we pick up a pair of sticks, the pair of sticks for one of the original n sticks." Let us first analyze A_1. The first time we pick up a stick we have $2n$ choices and once that choice has been made we must choose the complementary stick from among the $2n - 1$ remaining sticks. This results in a probability of

$$P(A_1) = \frac{2n}{2n(2n-1)} = \frac{1}{2n-1}.$$

Now we can more easily analyze the k^{th} step given that at the previous step we chose an original pair. At the k^{th} step, there are $2(n-k+1)$ (which consist of $n-k+1$ original pairs) remaining. Once we make a choice from among the $2(n-k+1)$ sticks, we must choose the complementary stick from among the $2(n-k+1)-1$ remaining sticks giving us

$$P\left(A_k \,\middle|\, \bigcap_{j=1}^{k-1} A_j\right) = \frac{2(n-k+1)}{(2(n-k+1))(2(n-k+1)-1)} = \frac{1}{2(n-k+1)-1}$$

By the hierarchical multiplicative formula,

$$P\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} P\left(A_k \,\middle|\, \bigcap_{j=1}^{k-1} A_j\right) = \left(\frac{1}{2n-1}\right) \cdots \left(\frac{1}{3}\right)\left(\frac{1}{1}\right).$$

For part (b): Let A_k denote the event "at the k^{th} time we pick up a pair of sticks, there is one short and one long stick." First, let us examine the probability of the first event in our sequence A_1. Initially there are $2n$ choices and once we make that choice (of either a long or a short stick) only n choices for the complementary stick. This gives us a probability of

$$P(A_1) = \frac{2nn}{2n(2n-1)} = \frac{n}{2n-1}.$$

A similar analysis to the one we provided above leads us to conclude that the probability at the k^{th} step of this process is given by

$$P(A_k) = \frac{2(n-k)(n-k)}{2(n-k)(2(n-k)-1)} = \frac{n-k}{2(n-k)-1}.$$

Once again, the hierarchical multiplicative formula gives us the probability we are after,

$$P\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} P\left(A_k \,\middle|\, \bigcap_{j=1}^{k-1} A_j\right) = \left(\frac{n}{2n-1}\right) \cdots \left(\frac{2}{3}\right)\left(\frac{1}{1}\right). \qquad \blacksquare$$

PROBLEM 1.3.2 (Handout 3, # 5). In a town, there are three plumbers. On a certain day, four residents need a plumber and they each call one plumber at random.

(a) What is the probability that all the calls go to one plumber (not necessarily a specific one)?
(b) What is the expected value of the number of plumbers who get a call?

SOLUTION. For part (a): There are 3^4 possible outcomes (since each of the four the four town residents have three choices of plumber whom to call). There are $\binom{3}{1} = 3$ ways to choose one plumber from among the three so the probability of A that the four residents all call the same plumber is

$$P(A) = \frac{3}{3^4} = \frac{1}{3^3} = \frac{1}{27} \approx 0.0370.$$

For part (b): Let $X\colon \Omega \to \mathbb{R}$ denote the number of plumbers that receive a call from one of the four residents. Before proceeding, we must find the probability $P(X = k)$ for $k = 2, 3$ (the case $k = 1$ was worked out in part (a)). The probability that two distinct plumbers are called can be broken into the probability that two of the residents call one plumber and the remaining residents call another plus the probability that three residents call one plumber and the remaining resident calls another; in symbols,

$$P(X = 2) = \frac{\binom{4}{2}3 \cdot 2}{3^4} + \frac{\binom{4}{3}3 \cdot 2}{3^4} = \frac{60}{81} \approx 0.7407.$$

Similarly, the probability that three distinct plumbers are called by the four residents is

$$P(X = 3) = \frac{3!3}{3^4} = \frac{6}{27} \approx 0.2222 = 1 - P(X = 2) - P(X = 1).$$

Therefore, the mean is

$$E(X) = 1 \cdot \tfrac{1}{27} + 2 \cdot \tfrac{20}{27} + 3 \cdot \tfrac{2}{9} \approx 2.1852.$$ ■

PROBLEM 1.3.3 (Handout 4, # 7 – *Polygraphs*). Polygraphs are routinely administered to job applicants for sensitive government positions. Suppose someone actually lying fails the polygraph 90% of the time. But someone telling the truth also fails the polygraph 15% of the time. If a polygraph indicates that an applicant is lying, what is the probability that he is in fact telling the truth? Assume a general prior probability p that the person is telling the truth.

SOLUTION. Let T denote the event that a given person is telling the truth and F denote the event that said person fails the polygraph test. Then Bayes' theorem implies that

$$\begin{aligned} P(T \mid F) &= \frac{P(F \mid T)P(T)}{P(F \mid T)P(T) + P(F \mid L)P(L)} \\ &= \frac{0.15p}{0.15p + 0.9 - 0.9p} \\ &= \frac{0.15p}{0.9 - 0.75p}. \end{aligned}$$ ■

PROBLEM 1.3.4 (Handout 4, # 8). In a bolt factory machines A, B, C manufacture, respectively, 25, 35, and 40 percent of the total. Of their output 5, 4, and 2 per cent are defective bolts. A bolt is drawn at random from the produce and is found defective. What are the probabilities that it was manufactured by machines A, B, C?

SOLUTION. Let denote the event that ■

PROBLEM 1.3.5 (Handout 4, # 9). Suppose that 5 men out of 100 and 25 women out of 10000 are colorblind. A colorblind person is chosen at random. What is the probability of his being male? (Assume males and females to be in equal numbers.)

SOLUTION. ■

PROBLEM 1.3.6 (Handout 4, # 10 – *Bridge*). In a Bridge party West has no ace. What probability should he attribute to the event of his partner having

(a) no ace;
(b) two or more aces?

Verify the result by a direct argument.

SOLUTION. ■

PROBLEM 1.3.7 (Handout 4, # 12). A true-false question will be posed to a couple on a game show. The husband and the wife each has a probability p of picking the correct answer. Should they decide to let one of the answer the question, or decide that they will give the common answer if they agree and toss a coin to pick the answer if they disagree?

SOLUTION. ■

PROBLEM 1.3.8 (Handout 4, # 13). An urn containing 5 balls has been filled up by taking 5 balls at random from a second urn which originally had 5 black and 5 white balls. A ball is chosen at random from the first urn and is found to be black. What is the probability of drawing a white ball if a second ball is chosen from among the remaining 4 balls in the first urn?

SOLUTION. ■

PROBLEM 1.3.9 (Handout 4, # 15). Events A, B, C have probabilities p_1, p_2, p_3. Given that exactly two of the three events occured, the probability that C occured is greater than $\frac{1}{2}$ if and only if ... (write down the necessary and sufficient condition).

SOLUTION. ■

Problem 1.3.10 (Handout 5, # 1). There are five coins on a desk: 2 are double-headed, 2 are double-tailed, and 1 is a normal coin.
One of the coins is selected at random and tossed. It shows heads.
What is the probability that the other side of this coin is a tail?

Solution. ∎

Problem 1.3.11 (Handout 5, # 2 – *Genetic testing*). There is a 50-50 chance that the Queen carries the gene for hemophilia. If she does, then each Prince has a 50-50 chance of carrying it. Three Princesses were recently tested and found to be non-carriers. Find the following probabilities:

(a) that the Queen is a carrier;
(b) that the fourth Princess is a carrier.

Solution. ∎

Problem 1.3.12 (Handout 5, # 4 – *Is Johnny in Jail*). Johnny and you are roommates. You are a terrific student and spend Friday evenings drowned in books. Johnny always goes out on Friday evenings. 40% of the times, he goes out with his girlfriend, and 60% of the times he goes to a bar. If he goes out with his girlfriend, 30% of the times he is just too lazy to come back and spends the night at hers. If he goes to a bar, 40% of the times he gets mad at the person sitting on his right, beats him up, and goes to jail.
On one Saturday morning, you wake up to see Johnny is missing. Where is Johnny?

Solution. ∎

1.4 Homework 4

PROBLEM 1.4.1 (Handout 5, # 2). In an urn, there are 12 balls. 4 of these are white. Three players: A, B, and C, take turns drawing a ball from the urn, in the alphabetical order. The first player to draw a white ball is the winner. Find the respective winning probabilities: assume that at each trial, the ball drawn in the trial before is put back into the urn (i.e., selection *with replacement*).

SOLUTION. ■

PROBLEM 1.4.2 (Handout 5, # 8). Consider n families with 4 children each. How large must n be to have a 90% probability that at least 3 of the n families are all girl families?

SOLUTION. ■

PROBLEM 1.4.3 (Handout 5, # 10 – *Yahtzee*). In Yahtzee, five fair dice are rolled. Find the probability of getting a Full House, which is three rolls of one number and two rolls of another, in Yahtzee.

SOLUTION. ■

PROBLEM 1.4.4 (Handout 5, # 12). The probability that a coin will show all heads or all tails when tossed four times is 0.25. What is the probability that it will show two heads and two tails?

SOLUTION. ■

PROBLEM 1.4.5 (Handout 5, # 13). Let the events $A_1, A_2, \ldots, A_n$ be independent and $P(A_k) = p_k$. Find the probability p that none of the events occurs.

SOLUTION. ■

PROBLEM 1.4.6 (Handout 6, # 5). Suppose a fair die is rolled twice and suppose X is the absolute value of the difference of the two rolls. Find the PMF and the CDF of X and plot the CDF. Find a median of X; is the median unique?

SOLUTION. ■

PROBLEM 1.4.7 (Handout 6, # 7). Find a discrete random variable X such that $E(X) = E(X^3) = 0$; $E(X^2) = E(X^4) = 1$.

SOLUTION. ■

Problem 1.4.8 (Handout 6, # 9 – *Runs*). Suppose a fair die is rolled n times. By using the indicator variable method, find the expected number of times that a six is followed by at least two other sixes. Now compute the value when $n = 100$.

Solution. ∎

Problem 1.4.9 (Handout 6, # 10 – *Birthdays*). For a group of n people find the expected number of days of the year which are birthdays of exactly k people. (Assume 365 days and that all arrangements are equally probable.)

Solution. ∎

Problem 1.4.10 (Handout 6, # 11 – *Continuation*). Find the expected number of multiple birthdays. How large should n be to make this expectation exceed 1?

Solution. ∎

Problem 1.4.11 (Handout 6, # 12 – *The blood-testing problem*). A large number, N, of people are subject to a blood test. This can be administered in two ways, (i) Each person can be tested separately. In this case N tests are required, (ii) The blood samples of k people can be pooled and analyzed together. If the test is negative, this one test suffices for the k people. If the test is positive, each of the k persons must be tested separately, and in all $k + 1$ tests are required for the k people. Assume the probability p that the test is positive is the same for all people and that people are stochastically independent.

(b) What is the expected value of the number, X, of tests necessary under plan (ii)?
(c) Find an equation for the value of k which will minimize the expected number of tests under the second plan. (Do not try numerical solutions.)

Solution. ∎

Problem 1.4.12 (Handout 6, # 13 – *Sample structure*). A population consists of r (classes whose sizes are in the proportion $p_1 : p_2 : \cdots : p_r$. A random sample of size n is taken with replacement. Find the expected number of classes not represented in the sample.

Solution. ∎

1.5 Homework 5

PROBLEM 1.5.1 (Handout 7, # 6(d, f)). Find the variance of the following random variables

(d) $X = \#$ of tosses of a fair coin necessary to obtain a head for the first time.

(f) $X = \#$ matches observed in random sitting of 4 husbands and their wives in opposite sides of a linear table.
This is an example of the *matching problem.*

SOLUTION. ■

PROBLEM 1.5.2 (Handout 7, # 8 – *Nonexistence of variance*). (a) Show that for a suitable positive constant c, the function $p(x) = \frac{c}{x^3}$, $x = 1, \ldots$, is a valid probability mass function (PMF).

(b) Show that in this case, the expectation of the underlying random variable exists, but the variance does not!

SOLUTION. ■

PROBLEM 1.5.3 (Handout 7, # 9). In a box, there are 2 black and 4 white balls. These are drawn out one by one at random (without replacement).

(a) Let X be the draw at which the first black ball comes out. Find the mean the variance of X.

(b) Let X be the draw at which the second black ball comes out. Find the meman the variance of X.

SOLUTION. ■

PROBLEM 1.5.4 (Handout 7, # 10). Suppose X has a *discrete uniform distribution* on the set $\{1, \ldots, N\}$.
Find formulas for the mean and the variance of X.

SOLUTION. ■

PROBLEM 1.5.5 (Handout 7, # 11 – *Be Original!*). Give an example of a random variable with mean 1 and variance 100.

SOLUTION. ■

PROBLEM 1.5.6 (Handout 7, # 13 – *Be Original!*). Suppose a random variable X has the property that its second and fourth moment are both 1.
What can you say about the nature of X?

SOLUTION. ■

PROBLEM 1.5.7 (Handout 7, # 14 – *Be Original!*). One of the following inequalities is true in general for all nonnegative random variables. Identify which one!

$$E(X)E(X^4) \geq E(X^2)E(X^3);$$
$$E(X)E(X^4) \leq E(X^2)E(X^2).$$

SOLUTION. ∎

PROBLEM 1.5.8 (Handout 7, # 15). Suppose X is the number of heads obtained in 4 tosses of a fair coin.

Find the expected value of the weird function

$$\ln\bigl(2 + \sin(\tfrac{\pi}{4}x)\bigr).$$

SOLUTION. ∎

PROBLEM 1.5.9 (Handout 7, # 16). In a sequence of Bernoulli trials let X be the length of the run (of either successes or failures) started by the first trial.

(a) Find the distribution of X, $E(X)$, $\mathrm{Var}(X)$.

SOLUTION. ∎

PROBLEM 1.5.10 (Handout 7, # 17). A man with n keys wants to open his door and tries the keys independently and at random. Find the mean and variance of the number of trials

(a) if unsuccessful keys are not eliminated from further selections;
(b) if they are.

(Assume that only one key fits the door. The exact distributions are given in II, 7, but are not required for the present problem.)

SOLUTION. ∎

1.6 Homework 6

PROBLEM 1.6.1 (Handout 8, # 2). Identify the parameters n and p for each of the following binomial distributions:

(a) # boys in a family with 5 children;
(b) # correct answers in a multiple choice test if each question has a 5 alternatives, there are 25 questions, and the student is making guesses at random.

SOLUTION. For part (a): The distribution is a $\mathrm{Bin}(5, 0.5)$.
For part (b): The distribution is a $\mathrm{Bin}(25, 0.2)$. ■

PROBLEM 1.6.2 (Handout 8, # 10). A newsboy purchases papers at 20 cents and sells them for 35 cents. He cannot return unsold papers. If the daily demand for papers is modeled as a $\mathrm{Bin}(50, 0.5)$ random variable, what is the optimum number of papers the newsboy should purchase?

SOLUTION. Let x denote the optimum number of newspapers to be bought and $X \sim \mathrm{Bin}(50, 0.5)$ the demand. Then the profit Y is given by

$$\begin{aligned} Y &= 35 \min\{X, x\} - 20x \\ &= 17.5((X + x) - |X - x|) - 20x \\ &= 17.5X - 2.5x - 17.5|X - x|. \end{aligned}$$

Therefore, the expectation of Y is

$$\begin{aligned} E(Y) &= 17.5E(X) - 2.5x - 17.5E(|X - x|) \\ &= 437.5 - 2.5x - \sum_{k=0}^{x} (x - k)\binom{50}{k} 0.5^{50} - \sum_{k=x+1}^{50} (k - x)\binom{50}{k} 0.5^{50}. \end{aligned}$$

Recursively trying values of x we find that the optimal value is $x = 24$. ■

PROBLEM 1.6.3 (Handout 8, # 12). How many independent bridge dealings are required in order for the probability of a preassigned player having four aces at least once to be $\frac{1}{2}$ or better? Solve again for some player instead of a given one.

SOLUTION. ■

PROBLEM 1.6.4 (Handout 8, # 13). A book of 500 pages contain s500 misprints. Estimate the chances that a given page contains at least three misprints.

SOLUTION. ■

PROBLEM 1.6.5 (Handout 8, # 14). Colorblindness appears in one percent of the people in a certain population. How large must a random sample (with replacements) be if the probability of its containing a colorblind person is to be 0.95 or more?

Solution. ■

Problem 1.6.6 (Handout 8, # 15). Two people toss a true coin n times each. Find the probability that they will score the same number of heads.

Solution. ■

Problem 1.6.7 (Handout 8, # 16 – *Binomial approximation to the hypergeometric distribution*). A population of TV elements is divided into red and black elements in the proportion $p : q$ (where $p + q = 1$). A sample of size n is taken without replacement. The probability that it contains exactly k red elements is given by the hypergeometric distribution of II, 6. Show that as $n \to \infty$ this probability approaches $\mathrm{Bin}(n, p)$.

Solution. ■

Problem 1.6.8 (Handout 9, # 3). Suppose X, Y, Z are mutually independent random variables, and $E(X) = 0$, $E(Y) = -1$, $E(Z) = 1$, $E(X^2) = 4$, $E(Y^2) = 3$, $E(Z^2) = 10$. Find the variance and the second moment of $2Z - \frac{Y}{2} + \mathrm{e}X$, where e is the number such that $\ln \mathrm{e} = 1$.

Solution. ■

Problem 1.6.9 (Handout 9, # 14 – *Variance of Product*). Suppose X, Y are independent random variables. Can it ever be true that $\mathrm{Var}(XY) = \mathrm{Var}(X)\,\mathrm{Var}(Y)$?
If it can, when?

Solution. ■

1.7 Homework 7

PROBLEM 1.7.1 (Handout 10, # 4 – *Poisson Approximation*). One hundred people will each toss a fair coin 200 times. Approximate the probability that at least 10 of the 100 people would each have obtained exactly 100 heads and 100 tails.

SOLUTION. ■

PROBLEM 1.7.2 (Handout 10, # 5 – *A Pretty Question*). Suppose X is a Poisson distributed random variable. Can three different values of X have an equal probability?

SOLUTION. ■

PROBLEM 1.7.3 (Handout 10, # 6 – *Poisson Approximation*). There are 20 couples seated at a rectangular table, husbands on one side and the wives on the other, in a random order. Using a Poisson approximation, find the probability that exactly two husbands are seated directly across from their wives; at least three are; at most three are.

SOLUTION. ■

PROBLEM 1.7.4 (Handout 10, # 7 – *Poisson Approximation*). There are 5 coins on a desk, with probabilities 0.05, 0.1, 0.05, 0.01, and 0.04 for heads. By using a Poisson approximation, find the probability of obtaining at least one head when the five coins are each tossed once. Is the number of heads obtained binomially distributed in this problem?

SOLUTION. ■

PROBLEM 1.7.5 (Handout 10, # 8). A book of 500 pages contains 500 misprints. Estimate the chances that a given page contains at least three misprints.

SOLUTION. ■

PROBLEM 1.7.6 (Handout 10, # 9). Estimate the number of raisins which a cookie should contain on the average if it is desired that not more than one cookie out of a hundred should be without raisin.

SOLUTION. ■

PROBLEM 1.7.7 (Handout 10, # 10). The terms $p(k;\lambda)$ of the Poisson distribution reach their maximum when k is the largest integer not exceeding λ.

SOLUTION. ■

Problem 1.7.8 (Handout 10, # 11). Prove

$$p(0;\lambda) + \cdots + p(n;\lambda) = \frac{1}{n!}\int_{\lambda}^{\infty} \mathrm{e}^{-x}x^n\,dx.$$

Solution. ∎

Problem 1.7.9 (Handout 10, # 12). There is a random number N of coins in your pocket, where N has a Poisson distribution with mean μ. Each one is tossed once.
Let X be the number of times a head shows.
Find the distribution of X.

Solution. ∎

Problem 1.7.10 (Handout 10, # 14). Find the MGF of a general Poisson distribution, and hence prove that the mean and the variance of an arbitrary Poisson distribution are equal.

Solution. ∎

Problem 1.7.11 (Handout 10, # 17 (a) – *Poisson approximations*). 20 couples are seated in a rectangular table, husbands on one side and the wives on the other. First, find the expected number of husbands that sit directly across from their wives. Then, using a Poisson approximation, find the probability that two do; three do; at most five do.

Solution. ∎

1.8 Homework 8

PROBLEM 1.8.1 (Handout 12, # 2). Let X be $U[a,b]$. Find the PDF, CDF, mean, and variance of X.

SOLUTION. ∎

PROBLEM 1.8.2 (Handout 12, # 8). The diameter of a circular disk cut out by a machine has the following PDF

$$f(x) = \begin{cases} \dfrac{4x - x^2}{9} & \text{for } 1 \leq x \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

Find the average diameter of disks coming from this machine (in inches).

SOLUTION. ∎

PROBLEM 1.8.3 (Handout 12, # 9). Suppose X is $U[0, 2\pi]$. Find $P(-0.5 \leq \sin X \leq 0.5)$.

SOLUTION. ∎

PROBLEM 1.8.4 (Handout 12, # 13). X has a piecewise uniform distribution on $[0,1]$, $[1,3]$, $[3,6]$, and $[6,10]$. Write its density function.

SOLUTION. ∎

PROBLEM 1.8.5 (Handout 12, # 16). Show that for every p, $0 \leq p \leq 1$, the function $f(x) = p \sin x + (1-p) \cos x$, $0 \leq x \leq \frac{\pi}{2}$ (and $f(x) = 0$ otherwise), is a density function. Find its CDF and use it to find all the medians.

SOLUTION. ∎

PROBLEM 1.8.6 (Handout 12, # 17). Give an example of a density function on $[0,1]$ by giving a formula such that the density is finite at zero, unbounded at one, has a unique minimum in the open interval $(0,1)$ and such that the median is 0.5.

SOLUTION. ∎

PROBLEM 1.8.7 (Handout 12, # 18 – *A Mixed Distribution*). Suppose the damage claims on a particular type of insurance policy are uniformly distributed on $[0,5]$ (in thousands of dollars), but the maximum by the insurance company is 2500 dollars. Find the CDF and the expected value of the payout, and plot the CDF. What is unusual about this CDF?

Solution. ■

Problem 1.8.8 (Handout 12, # 19 – *Random Distribution*). Jen's dog broke her six-inch long pencil off at a random point on the pencil. Find the density function and the expected value of the ratio of the lengths of the shorter piece and the longer piece of the pencil.

Solution. ■

Problem 1.8.9 (Handout 12, # 20 – *Square of a PDF Need Not Be a PDF*). Give an example of a density function $f(x)$ on $[0, 1]$ such that $cf^2(x)$ cannot be a density function for any c.

Solution. ■

Problem 1.8.10 (Handout 12, # 21 – *Percentiles of the Standard Cauchy*). Find the p^{th} percentile of the standard Cauchy density for a general p, and compute it for $p = 0.75$.

Solution. ■

Problem 1.8.11 (Handout 12, # 22 – *Integer Part*). Suppose X has a uniform distribution on $[0, 10.5]$. Find the expected value of the integer part of X.

Solution. ■

Problem 1.8.12 (Handout 12, # 23). X is uniformly distributed on some interval $[a, b]$. If its mean is 2, and variance is 3, what are the values of a, b?

Solution. ■

1.9 Homework 9

PROBLEM 1.9.1 (Handout 13, # 7). Let X have a *double exponential* density $f(x) = \frac{1}{2\sigma}\mathrm{e}^{-\frac{|x|}{\sigma}}$, $-\infty < x < \infty$, $\sigma > 0$.

(a) Show that all moments exist for this distribution.
(b) However, show that the MGF exists only for restricted values. Identify them and find a formula.

SOLUTION. ∎

PROBLEM 1.9.2 (Handout 13, # 16). Give an example of each of the following phenomena:

(a) A continuous random variable taking values in $[0, 1]$ with equal mean and median.
(b) A continuous random variable taking values in $[0, 1]$ with mean equal to twice the median.
(c) A continuous random variable for which the mean does not exist.
(d) A continuous random variable for which the mean exists, but the variance does not exist.
(e) A continuous random variable with a PDF that is not differentiable at zero.
(f) a positive continuous random variable for which the mode is zero, but the mean does not exist.
(g) A continuous random variable for which all moments exist.
(h) A continuous random variable with median equal to zero, and 25^{th} and 75^{th} percentiles equal to 1.
(i) A continuous random variable X with mean equal to median equal to mode equal to zero, and $E(\sin X) = 0$.

SOLUTION. ∎

PROBLEM 1.9.3 (Handout 13, # 17). An exponential random variable with mean 4 is known to be larger than 6. What is the probability that it is larger than 8?

SOLUTION. ∎

PROBLEM 1.9.4 (Handout 13, # 18 – *Sum of Gammas*). Suppose X, Y are independent random variables, and $X \sim G(\alpha, \lambda)$, $Y \sim G(\beta, \lambda)$. Find the distribution of $X+Y$ by using moment-generating functions.

SOLUTION. ∎

PROBLEM 1.9.5 (Handout 13, # 19 – *Product of Chi Squares*). Suppose $X_1, X_2, \ldots, X_n$ are independent chi square variables, with $X_k \sim \chi^2_{m_k}$. Find the mean and variance of $\prod_{k=1}^{n} X_k$.

SOLUTION. ∎

Problem 1.9.6 (Handout 13, # 20). Let $Z \sim N(0,1)$. Find

$$P(0.5 < |Z - 0.5| < 1.5); \quad P\left(\frac{\mathrm{e}^Z}{1+\mathrm{e}^Z} > \frac{3}{4}\right); \quad P(\Phi(Z) < 0.5).$$

Solution. ■

Problem 1.9.7 (Handout 13, # 21). Let $Z \sim N(0,1)$. Find the density of $\frac{1}{Z}$. Is the density bounded?

Solution. ■

Problem 1.9.8 (Handout 13, # 22). The 25^{th} and the 75^{th} percentile of a normally distributed random variable are -1 and 1. What is the probability that the random variable is between -2 and 2?

Solution. ■

1.10 Homework 10

PROBLEM 1.10.1 (Handout 14, # 5). Approximately find the probability of getting a total exceeding 3600 in 1000 rolls of a fair die.

SOLUTION. ■

PROBLEM 1.10.2 (Handout 14, # 6). A basketball player has a history of converting 80% of his free throws. Find a normal approximation with a continuity correction of the probability that he will make between 18 and 22 throws out of 25 free throws.

SOLUTION. ■

PROBLEM 1.10.3 (Handout 14, # 7). Suppose $X_1, \ldots, X_n$ are independent $N(0, 1)$ variables. Find an approximation to the probability that $\sum_{k=1}^{n} X_k$ is larger than $\sum_{k=1}^{n} X_k^2$, when $n = 10, 20, 30$.

SOLUTION. ■

PROBLEM 1.10.4 (Handout 14, # 8 – *A Product Problem*). Suppose $X_1, \ldots, X_{30}$ are 30 independent variables, each distributed as $U[0, 1]$. Find an approximation to the probability that their *geometric mean* exceeds 0.4; exceeds 0.5.

SOLUTION. ■

PROBLEM 1.10.5 (Handout 14, # 9 – *Comparing a Poisson Approximation and a Normal Approximation*). Suppose 1.5% of residents of a town never read a newspaper. Compute the exact value, a Poisson approximation, and a normal approximation of the probability that at least one resident in a sample of 50 residents never reads a newspaper.

SOLUTION. ■

PROBLEM 1.10.6 (Handout 14, # 10 – *Test Your Intuition*). Suppose a fair coin is tossed 100 times. Which is more likely: you will get exactly 50 heads, or you will get more than 60 heads?

SOLUTION. ■

PROBLEM 1.10.7 (Handout 14, # 11). Find the probability that among 10000 random digits the digit 7 appears not more than 968 times.

SOLUTION. ■

PROBLEM 1.10.8 (Handout 14, # 12). Find a number k such that the probability is about 0.5 that the number of heads obtained in 1000 tossings of a coin will be between 490 and k.

SOLUTION. ∎

PROBLEM 1.10.9 (Handout 14, # 13). In 10000 tossings, a coin fell heads 5400 times. Is it reasonable to assume that the coin is skew?

SOLUTION. ∎

PROBLEM 1.10.10 (Handout 14, # 14). Interpret in plain words the statement the problem: *(Normal approximation to the Poisson distribution).* Using Stirling's formula, show that, if $\lambda \to \infty$, then for fixed $\alpha < \beta$

$$\sum_{\lambda+\alpha\sqrt{\lambda}<k<\lambda+\beta\sqrt{\lambda}} p(k;\lambda) \longrightarrow \Phi(\beta) - \Phi(\alpha).$$

SOLUTION. Recall that $p(k;\lambda)$ is the discrete Poisson distribution

$$p(k;\lambda) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}.$$

∎

PROBLEM 1.10.11 (Handout 14, # 15). Give a proof that as $x \to \infty$,

$$1 - \Phi(x) \sim \frac{\phi(x)}{x}.$$

Remark: This gives the exact rate at which the standard normal right tail area goes to zero. It is even faster than the rate at which the standard normal density goes to zero, because of the extra x in the denominator.

SOLUTION. ∎

1.11 Homework 11

PROBLEM 1.11.1 (DasGupta 7.2 (a), (b), (c), (d), (e)). (a) Suppose $E|X_n - c|^\alpha \to 0$, where $0 < \alpha < 1$. Does X_n necessarily converge in probability to c?

(b) Suppose $a_n(X_n - \theta) \xrightarrow{\mathcal{L}} N(0,1)$. Under what condition on a_n can we conclude that $X_n \xrightarrow{\mathcal{P}} \theta$?

(c) $o_p(1) + O_p(1) =?$

(d) $o_p(1)O_p(1) =?$

(e) $o_p(1) + o_p(1)O_p(1) =?$

SOLUTION. ∎

PROBLEM 1.11.2 (DasGupta 7.3 [Monte Carlo]). Consider the purely mathematical problem of finding a definite integral $f(x)\,dx$ for some (possibly complicated) function $f(x)$. Show that the SLLN provides a method for approximately finding the value of the integral by using appropriate averages $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$.

Numerical analysts call this Monte Carlo integration.

SOLUTION. ∎

PROBLEM 1.11.3 (DasGupta 7.4 (a), (b)). Suppose $X_1, \ldots,$ are i.i.d. and that $E(X_1) = \mu \neq 0$, $\mathrm{Var}(X_1) = \sigma^2 < \infty$. Let $S_{m,p} = \sum_{k=1}^{m} X_k^p$, $m \geq 1$, $p = 1, 2$.

(a) Identify with proof the almost sure limit of $S_{m,1}/S_{n,1}$ for fixed m, and $n \to \infty$.

(b) Identify with proof the almost sure limit of $S_{n-m,1}/S_{n,1}$ for fixed m, and $n \to \infty$.

SOLUTION. ∎

PROBLEM 1.11.4 (DasGupta 7.5 (a)). Let A_n, $n \geq 1$, A be events with respect to a common sample space Ω.

(a) Prove that $I_{A_n} \xrightarrow{\mathcal{L}} I_A$ if and only if $P(A_n) \to P(A)$.

SOLUTION. ∎

PROBLEM 1.11.5 (DasGupta 7.11 [Sample Maximum]). Let X_k, $k \geq 1$, be an i.i.d. sequence, and $X_{(n)}$ the maximum of $X_1, \ldots, X_n$. Let $\xi(F) = \sup\{\, x : F(x) < 1 \,\}$, where F is the common CDF of the X_k. Prove that $X_{(n)} \xrightarrow{\text{a.s.}} \xi(F)$.

SOLUTION. ∎

PROBLEM 1.11.6 (DasGupta 7.14 (a)). Suppose X_k are i.i.d. standard Cauchy. Show that

(a) $P(|X_n| > n \text{ infinitely often}) = 1$.

Solution. ■

Problem 1.11.7 (DasGupta 7.16 [Coupon Collection]). Cereal boxes contain independently and with equal probability exactly one of n different celebrity pictures. Someone having the entire set of n pictures can cash them in for money. Let W_n be the minimum number of cereal boxes one would need to purchase to own a complete set of the pictures. Find a sequence a_n such that $\frac{W_n}{a_n} \xrightarrow{\mathcal{P}} 1$. *Hint:* Approximate the mean of W_n.

Solution. ■

Problem 1.11.8 (DasGupta 7.17). Let $X \sim \text{Bin}(n, p)$. Show that $(\frac{X_n}{n})^2$ and $\frac{X_n(X_n-1)}{n(n-1)}$ both converging in probability to p^2. Do they converge almost surely?

Solution. ■

Problem 1.11.9 (DasGupta 7.21). Let $X_1, X_2, \ldots,$ be i.i.d. $U[0,1]$. Let

$$G_n = (X_1 \cdots X_n)^{1/n}.$$

Find c such that $G_n \xrightarrow{\mathcal{P}} c$.

Solution. ■

Problem 1.11.10 (DasGupta 7.30 [Conceptual]). Suppose $X_n \xrightarrow{\mathcal{L}} X$, and also $Y_n \xrightarrow{\mathcal{L}} X$. Does this mean that $X_n - Y_n$ converge in distribution to (the point mass at) zero?

Solution. ■

Problem 1.11.11 (DasGupta 7.31 (a)). (a) Suppose $a_n(X_n - \theta) \to N(0, \tau^2)$; what can be said about the limiting distribution of $|X_n|$, when $\theta \neq 0$, $\theta = 0$?

Solution. ■

1.12 Homework 12

PROBLEM 1.12.1 (Handout 15, # 10). Consider the experiment of picking one word at random from the sentence

> *All is well in the newell family*

Let X be the length of the word selected and Y the number of Ls in it. Find in a tabular form the joint PMF of (X, Y), their marginal PMFs, means, and variances, and the correlation between X and Y.

SOLUTION. The joint PMF of (X, Y) is given by

$Y \backslash X$	2	3	4	5	6
0	$\frac{2}{7}$	$\frac{1}{7}$	0	0	0
1	0	0	0	0	$\frac{1}{7}$
2	0	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	0

The marginal PMF of X is thus given by

$$f_X(x) = \begin{cases} \frac{2}{7} & \text{for } x = 2, 3, \\ \frac{1}{7} & \text{for } x = 4, 5, 6 \end{cases}$$

and the marginal PMF of Y is given by

$$f_Y(x) = \begin{cases} \frac{3}{7} & \text{for } x = 0, 2, \\ \frac{1}{7} & \text{for } x = 1. \end{cases}$$

So the mean and variance of X and Y are

$$\begin{aligned} \mu_X &= \frac{4+6+4+5+6}{7} & \mu_Y &= 1, \\ &= \frac{25}{7}, \\ \operatorname{Var}(X) &= \frac{8+18+16+25+36}{7} - \left(\frac{25}{7}\right) & \operatorname{Var}(Y) &= \frac{1+12}{7} - 1 \\ &= \frac{96}{49}, & &= \frac{6}{7}. \end{aligned}$$

Lastly, the correlation between X and Y is

$$\rho_{X,Y} = \frac{5}{\sqrt{\frac{576}{7}}} \approx 0.5512. \qquad \blacksquare$$

PROBLEM 1.12.2 (Handout 15, # 11). Consider the joint PMF $p(x, y) = cxy$, $1 \leq x \leq 3$, $1 \leq y \leq 3$.

(a) Find the normalizing constant c.
(b) Are X and Y independent? Prove your claim.

(c) Find the expectations of X, Y, and XY.

Solution. *Remark:* Note that below parts (a), (b), and (c) are out of order.

For part (a): The normalizing constant is $c = \frac{1}{36}$; this is because

$$\sum_{x,y=(1,1)}^{(3,3)} cxy = 36c$$

For part (c): First,

$$E(X) = E(Y) = \sum_{x=1}^{3} x^2(1+2+3)c = 6c\sum_{x=1}^{3} x^2 = \frac{7}{3}$$

and

$$E(XY) = \sum_{(x,y)=(1,1)}^{(3,3)} cx^2y^2 = \frac{49}{9}$$

For part (b): We see that X and Y are independent; $E(XY) = E(X)E(Y)$. ∎

Problem 1.12.3 (Handout 15, # 12). A fair die is rolled twice. Let X be the maximum and Y the minimum of the two rolls. By using the joint PMF of X and Y worked out in the text, find the PMF of $\frac{X}{Y}$, and hence the mean of $\frac{X}{Y}$.

Solution. The PMF of $\frac{X}{Y}$ is given by

$$f_{\frac{X}{Y}}(x) = \begin{cases} \frac{1}{6} & \text{for } x = 1, 2, \\ \frac{1}{9} & \text{for } x = \frac{3}{2}, 3, \\ \frac{1}{18} & \text{for} x = \frac{5}{2}, 4, 5, 6, \frac{5}{3}, \frac{4}{3}, \frac{5}{4}, \frac{5}{6}. \end{cases}$$

So that the mean is

$$\mu_{\frac{X}{Y}} = \frac{487}{216} \approx 2.2546.$$ ∎

Problem 1.12.4 (Handout 15, # 13). Two random variables have the joint PMF $p(x, x+1) = \frac{1}{n+1}$, $x = 0, \dots, n$. Answer the following question with as little calculation as possible.

(a) Are X and Y independent?
(b) What is the variance of $Y - X$?
(c) What is $\mathrm{Var}(Y \mid X = 1)$?

Solution. For part (a): No. The probability that $Y = 2$ given that $X = 1$ is 1, but the probability that $Y = 2$ is $\frac{1}{n+1}$.

For part (b): $\mathrm{Var}(Y - X) = 0$, because $Y - X$ is constant; it is always 1.

For part (c): $\mathrm{Var}(Y \mid X = 1) = 0$, because $Y = 2$ if $X = 1$. ∎

Problem 1.12.5 (Handout 15, # 14 – *Binomial Conditional Distribution*). Suppose X and Y are independent random variables, and $X \sim \text{Bin}(m, p)$, $Y \sim \text{Bin}(n, p)$. Show that the conditional distribution of X given by $X + Y = t$ is a hypergeometric distribution; identify the parameters of this hypergeometric distribution.

Solution. First, let us find the PMF of X given $X + Y = t$:

$$\begin{aligned}
P(X = x \mid X + Y = t) &= \frac{P\big(\{ X = x \} \cap \{ X + Y = t \}\big)}{P(X + Y = t)} \\
&= \frac{P(Y = t - x)}{P(X + Y = t)} \\
&= \frac{\binom{n}{x}\binom{m}{t-x} p^t (1-p)^{m+n-t}}{\binom{m+n}{t} p^t (1-p)^{m+n-t}} \\
&= \frac{\binom{n}{x}\binom{m}{t-x}}{\binom{m+n}{t}}.
\end{aligned}$$

This distribution is precisely $\text{Hypergeo}(t, m, n + m)$. ∎

Problem 1.12.6 (Handout 15, # 15). Suppose a fair die is rolled twice. Let X and Y be the two rolls. Find the following with as little calculation as possible.

(a) $E(X + Y \mid Y = y)$.
(b) $E(XY \mid Y = y)$.
(c) $\text{Var}(X^2 Y \mid Y = y)$.
(d) $\rho_{X+Y,X-Y}$.

Solution. For part (a):

$$E(X + Y \mid Y = y) = E(X \mid Y = y) + E(Y \mid Y = y) = 3.5 + y.$$

For part (b):

$$E(XY \mid Y = y) = E(X \mid Y = y)E(Y \mid Y = y) = 3.5y.$$

For part (c):

$$\text{Var}(X^2 Y \mid Y = y) = E((X^2 Y)^2 \mid Y = y) - E(X^2 Y \mid Y = y)^2 = c^2 \left(\frac{91}{6} - 3.5 \right).$$

For part (d):

$$\begin{aligned}
\text{Cov}(X + Y, X - Y) &= E((X + Y)(X - Y)) - E(X + Y)E(X - Y) \\
&= E(X)E(X) - E(Y)E(Y) - E(X)E(X) + E(Y)E(Y) \\
&= 0,
\end{aligned}$$

so $\rho_{X+Y,X-Y} = 0$. ∎

Problem 1.12.7 (Handout 15, # 16 – *A Standard Deviation Inequality*). Let X and Y be two random variables. Show that $\sigma_{X+Y} \le \sigma_X + \sigma_Y$.

Solution. Suppose σ_X and σ_Y exist and are finite. We want to show

$$\sigma_{X+Y} \le \sigma_X + \sigma_Y;$$

this is the same as showing that

$$\begin{gathered}\sigma_{X+Y}^2 \le \sigma_X + \sigma_Y^2 + 2\sigma_X\sigma_Y \\ \operatorname{Var}(X+Y) \le \operatorname{Var}(X) + \operatorname{Var}(Y) + 2[\operatorname{Var}(X)\operatorname{Var}(Y)]^{1/2}.\end{gathered}$$

First, let us expand $\operatorname{Var}(X+Y)$ using the definition of variance, we have

$$\begin{aligned}\operatorname{Var}(X+Y) &= E\big((X+Y)^2\big) - E(X+Y)^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 \\ &= \big(E(X^2) - E(X)^2\big) + \big(E(Y^2) - E(Y)^2\big) + 2[E(XY) - E(X)E(Y)] \\ &= \operatorname{Var}(X) + \operatorname{Var}(Y) + 2[E(XY) - E(X)E(Y)].\end{aligned}$$

Therefore, it suffices to show that

$$E(XY) - E(X)E(Y) \le [\operatorname{Var}(X)\operatorname{Var}(Y)]^{1/2},$$

or, rewritten using covariance,

$$\operatorname{Cov}(X,Y)^2 \le \operatorname{Var}(X)\operatorname{Var}(Y).$$

By the Cauchy–Schwartz inequality, we have

$$\begin{aligned}\operatorname{Cov}(X,Y)^2 &= E\big[(X - E(X))(Y - E(Y))\big]^2 \\ &\le E\big[(X - E(X))^2\big]E\big[(Y - E(Y))^2\big] \\ &= \operatorname{Var}(X)\operatorname{Var}(Y).\end{aligned}$$

■

Problem 1.12.8 (Handout 15, # 17). Seven balls are distributed randomly in seven cells. Let X_k be the number of cells containing exactly k balls. Using the probabilities tabulated in II, 5, write down the joint distribution of X_2, X_3.

Solution. The tabled referenced in this problem is on p. 40 of Feller. Let us write down a table of our own for the joint distribution of (X_2, X_3):

$X_3 \backslash X_2$	0	1	2	3
0	0.0475	0.1564	0.3213	0.1071
1	0.1089	0.2142	0.0268	0
2	0.0179	0	0	0

.

Let us do a sanity check by summing over all of the entries in the table above

$$0.0475 + 0.1564 + 0.3213 + 0.1071 + 0.1089 + 0.2142 + 0.0268 + 0 + 0.0179 + 0 + 0 + 0 \approx 1.$$

■

PROBLEM 1.12.9 (Handout 15, # 18). Two ideal dice are thrown. Let X be the score on the first die and Y be the larger of two scores.

(a) Write down the joint distribution of X and Y.
(b) Find the means, the variances, and the covariance.

SOLUTION. For part (a): The random variable X takes on integer values between zero and six and so does Y. Moreover, the dependence of Y on X tells us that $P(\{X = k\} \cap \{Y = l\}) = 0$ if $l < k$; this allows us to fill in a significant portion of the joint distribution table:

$Y \backslash X$	1	2	3	4	5	6
1	$\frac{1}{36}$	0	0	0	0	0
2	$\frac{1}{36}$	$\frac{2}{36}$	0	0	0	0
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{3}{36}$	0	0	0
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{4}{36}$	0	0
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	0
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$

(One can easily verify that the sum of the entries in this table do in fact add up to one.)

For part (b): We can recover the individual PMFs for X and Y using the table in part (a) and so recover the mean and variance. These are

$$\begin{aligned}
E(X) &= \frac{6}{36} + 2\left(\frac{6}{36}\right) + 3\left(\frac{6}{36}\right) + 4\left(\frac{6}{36}\right) + 5\left(\frac{6}{36}\right) + 6\left(\frac{6}{36}\right) \\
&= 3.5, \\
E(X) &= 1^2\left(\frac{6}{36}\right) + 2^2\left(\frac{6}{36}\right) + 3^2\left(\frac{6}{36}\right) + 4^2\left(\frac{6}{36}\right) + 5^2\left(\frac{6}{36}\right) + 6^2\left(\frac{6}{36}\right) \\
&\approx 15.1667, \\
\mathrm{Var}(X) &\approx 2.9167,
\end{aligned}$$

and

$$\begin{aligned}
E(Y) &= \frac{1}{36} + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right) \\
&\approx 4.4722, \\
E(Y^2) &= 1^2\left(\frac{1}{36}\right) + 2^2\left(\frac{3}{36}\right) + 3^2\left(\frac{5}{36}\right) + 4^2\left(\frac{7}{36}\right) + 5^2\left(\frac{9}{36}\right) + 6^2\left(\frac{11}{36}\right) \\
&\approx 21.9722, \\
\mathrm{Var}(Y) &\approx 1.9715,
\end{aligned}$$

and lastly (after a long calculation which we omit) the covariance is

$$\mathrm{Cov}(X, Y) \approx 2.06111. \qquad \blacksquare$$

PROBLEM 1.12.10 (Handout 15, # 19). Let X_1 and X_2 be independent and have the common geometric distribution $\{q^k p\}$ (as in problem 4). Show without calculations that the *conditional* *distribution of* X_1 *given* $X_1 + X_2$ *is uniform*, that is,

$$P(X_1 = k \mid X_1 + X_2 = n) = \frac{1}{n+1}, \quad k = 0, \dots, n. \tag{1.12.1}$$

Solution. By definition of conditional probability, we have

$$\begin{aligned} P(X_1 = k \mid X_1 + X_2 = n) &= \frac{P(\{\, X_1 = k \,\} \cap \{\, X_1 + X_2 = n \,\})}{P(X_1 = k)} \\ &= \frac{P(X_2 = n - k)}{P(X_1 + X_2 = n)} \\ &= \frac{q^{n-k} p}{q^{n-k} p (n+1)} \\ &= \frac{1}{n+1}. \end{aligned}$$

■

Problem 1.12.11 (Handout 15, # 20). If two random variables X and Y assume only two values each, and if $\operatorname{Cov}(X, Y) = 0$, then X and Y are independent.

Solution. We show that the joint PDF of (X, Y) is

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

Suppose X assumes the values $\{a, b\}$ and Y assumes the values $\{c, d\}$ where, without loss of generality, we may assume $a < b$ and $c < d$; however, we may have $a = c$, $b = c$, $a = d$, etc. Let p_a, p_b, p_c, and p_d be the probabilities associated to a, b, c, and d, respectively. Then, we have

$$p_a + p_b = 1, \quad p_c + p_d = 1,$$

and more significantly

$$\begin{aligned} \operatorname{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\ E(XY) &= (ap_a + bp_b)(cp_c + dp_d) \\ \sum_{\substack{x \in \{a,b\}, \\ y \in \{c,d\}}} xy f_{X,Y}(x, y) &= (ap_a + bp_b)(cp_c + dp_d) \\ \begin{array}{r} ac f_{X,Y}(a, c) + ad f_{X,Y}(a, d) \\ + bc f_{X,Y}(b, c) + bd f_{X,Y}(b, d) \end{array} &= \begin{array}{l} acp_a p_c + adp_a p_d \\ \quad + bcp_b p_c + bdp_b p_d. \end{array} \end{aligned}$$

A term by term comparison shows that we must have

$$f(x, y) = xyp_x p_y$$

for $x \in \{a, b\}$, $y \in \{c, d\}$. Thus, $f_{X,Y}(x, y) = f_X(x) f_Y(y)$; i.e., X and Y are independent. ■

1.13 Homework 13

PROBLEM 1.13.1 (Handout 17, # 16). Suppose $X \sim \mathrm{Exp}(1)$, $Y \sim U[0,1]$, and X, Y are independent.

(a)

SOLUTION. ■

PROBLEM 1.13.2 (Handout 17, # 18). Two points A, B are chosen at random from the unit circle. Find the probability that the circle centered at A with radius AB is fully contained within the original unit circle.

SOLUTION. ■

PROBLEM 1.13.3 (Handout 17, # 19). Let X, Y be i.i.d. $U[0,1]$ random variables. Find the correlation between $\max\{X, Y\}$ and $\min\{X, Y\}$.

SOLUTION. ■

1.14 Homework 14

PROBLEM 1.14.1 (Handout 18, # 15). (X, Y) is distributed uniformly inside of the unit circle. Find the density of $X + Y$ and hence the mean of $X + Y$. Was the value of the mean obvious? Why?

SOLUTION. ■

PROBLEM 1.14.2 (Handout 18, # 16). Let X be a random number in $[0, 1]$. What is the probability that the number 5 is completely missing from the decimal expansion of X?

SOLUTION. ■

PROBLEM 1.14.3 (Handout 18, # 17). A foot long stick is broken into three pieces. Find the density functions of the length of the longest part, the smallest part, and the medium part. What are the expected values for each part?

SOLUTION. Consider the intersection of the plane $X + Y + Z = 1$ with the region $\{ X, Y, Z \geq 0 \}$. ■

2 Midterms and Finals

2.1 Midterm 1

Problem 2.1.1.

Solution. ■

Problem 2.1.2.

Solution. ■

Problem 2.1.3.

Solution. ■

Problem 2.1.4.

Solution. ■

Problem 2.1.5.

Solution. ■

Problem 2.1.6.

Solution. ■

Problem 2.1.7.

Solution. ■

Problem 2.1.8.

Solution. ■

2.2 Midterm 2

Problem 2.2.1.

Solution. ∎

Problem 2.2.2.

Solution. ∎

Problem 2.2.3.

Solution. ∎

Problem 2.2.4.

Solution. ∎

Problem 2.2.5.

Solution. ∎

Problem 2.2.6.

Solution. ∎

Problem 2.2.7.

Solution. ∎

Problem 2.2.8.

Solution. ∎

3 Homework Solutions (Yip)

These are the solutions to Yip's Math/Stat 519 homework for the fall semester of 2016.
Our main reference is [3].

3.1 Homework 1

3.1.1 Problems

Problem 3.1.1 (Ross, §1, # 7).

(a) In how many different ways can 3 boys and 3 girls sit in a row?

(b) In how many ways can 3 boys and 3 girls sit in a row if the boys and the girls are each to sit together?

(c) In how many ways if only the boys must sit together?

(d) In how many ways if no two people of the same sex are allowed to sit together?

Solution. For the problems that follow, we will not justify our answer but instead we make a note here that they are consequences of the general principle of counting introduced in the book.

For part (a), there are 6! ways of arranging 3 boys and 3 girls (6 children).

For part (b), assuming the author meant "there is at least one girl next to every boy" by "the boys and girls must sit together" we have $2(3!)^2$ ways to arrange the 3 boys and 3 girls in this way (the 2 comes from the choice of placing a boy or a girl first).

For part (c), assuming the author means that "it does not matter in which order the boys sit, but they must sit together" we have 4(3!) where the 3! comes from the different ways we can arrange the 3 girls and the 4 from the different ways we can place 3 boys contiguously on a row with 6 slots.

For part (d), we can either start with a boy or a girl, which gives us a factor of 2, and once that choice has been made, the rest of the spots are reserved for boys/girls alternating from one to the other to the end of the row; that is, they are completely determined except for the particular boy or girl that is to take those spots. Next for each spot, we have 3 choices of boy (respectively, girl) which gives us a 3! factor. Thus, the number of ways in which we can do this is precisely $2(3!)^2$. ■

Problem 3.1.2 (Ross, §1, # 11). In how many ways can 3 novels, 2 mathematics books, and 1 chemistry book be arranged on a bookshelf if

(a) the books can be arranged in any order?

(b) the mathematics books must be together and the novels must be together?

(c) the novels must be together, but the other books can be arranged in any order?

Solution. For part (a), if the $3+2+1=6$ books can be arranged in any order, there are $6!$ arrangements.

For part (b), let us first count the different ways we can arrange 3 blocks of books. There are $3!$ different ways to arrange these and once this arrangement has been made, $3!2!$ ways to arrange the 3 novels and 2 mathematics books. Thus, there are $(3!)^2 2!$ ways of arranging the 6 books subject to these restrictions.

For part (c), if the novels must be together, we must first count the different number of ways we can put a contiguous row of 3 books in 6 slots. The answer to that problem is 4. Now, assuming the arrangement of the 3 novels does not matter, we have $4 \cdot 3!$ ways of arranging the rest of the books. If we care about randomizing the 3 novels themselves, then we have $4(3!)^2$ distinct arrangements. ∎

Problem 3.1.3 (Ross, §1, # 19). From a group of 8 women and 6 men, a committee consisting of 3 men and 3 women is to be formed. How many different committees are possible if

(a) 2 of the men refuse to serve together?

(b) 2 of the women refuse to serve together?

(c) 1 man and 1 woman refuse to serve together?

Solution. For part (a), there are

$$\binom{8}{3}\left[\binom{4}{3}+2\binom{4}{2}\right]$$

ways to choose a committee that avoids putting the two men in question together. The $\binom{8}{3}$ comes from the different number of ways we can choose 3 women from a pool of 8 and the $\binom{4}{3}+2\binom{4}{2}$ are the different number of ways we can choose 3 men to from a pool of 6, 2 of which refuse to serve together; i.e., we can either choose neither of the two men in question, giving us $\binom{4}{3}$ possibilities, or choose one of them giving us $\binom{4}{2}$ possibilities (multiply this by 2, this represents the choice for one of the 2 men).

For part (b), the exact argument given above yields

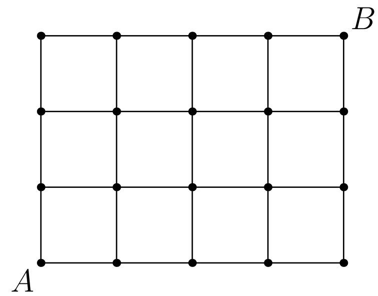
$$\binom{6}{3}\left[\binom{6}{3}+2\binom{6}{2}\right]$$

possible committees.

For part (c), we have $\binom{5}{3}\binom{7}{3}$ if neither the man nor the women in question serve in the committee; $\binom{5}{3}\binom{7}{2}$ if the woman serves; and $\binom{5}{2}\binom{7}{3}$ if the man serves. In total, this is

$$\binom{5}{3}\binom{7}{3}+\binom{5}{3}\binom{7}{2}+\binom{5}{2}\binom{7}{3}.$$ ∎

Problem 3.1.4 (Ross, §1, # 21). Consider the grid of points show at the top of the next column (in the book; we draw it here using Asymptote).



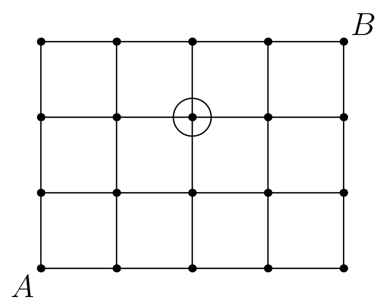
Suppose that, starting at the point labeled A, you can go one step up or one step to the right at each move. This procedure is continued until the point labeled B is reached. How many different paths from A to B are possible?
Hint: Note that to reach B from A you must take 4 steps to the right and 3 steps up.

Solution. There are $4 + 3 = 7$ total steps we must make to reach point B from point A. Of these steps 4 must be in the direction right and 3 must be in the direction left. Thus, by some formula in Ross's book, we have

$$\binom{7}{4,3}$$

total paths from point A to point B in the figure. ■

Problem 3.1.5 (Ross, §1, # 22). In Problem # 21, how many different paths are there from A to B that go through the point circled in the following lattice?



Solution. This problem is only slightly more complicated than its predecessor. We say slightly because thinking of the circled point as another point C and thinking of paths from A to B passing to C as joining paths from A to C and paths from C to D, we have $\binom{4}{2,2}$ ways to get from A to C and $\binom{3}{2,1}$ ways to get from C to B. Thus, we have

$$\binom{4}{2,2}\binom{3}{2,1}$$

paths from A to B passing through C. ■

Problem 3.1.6 (Ross, §1, # 33). We have \$20,000 that must be invested among 4 possible opportunities. Each investment must be integral in units of \$1000, and there are minimal investment that need to be made if one is to invest these opportunities. The minimal investments are \$2000, \$2000, \$3000, and \$4000. How many different investment strategies are available if

(a) an investment must be made in each opportunity?

(b) investments must be made in at least 3 of the 4 opportunities?

Solution. Let us first cut the complexity of the problem by changing translating from dollars (\$) to units (1 unit is \$1000). Thus, we have 20 units and we want to distribute them in integers between the 4 opportunities which require minimal investments of 2, 2, 3, and 4 units.

For part (a), if we are to invest in each opportunity we are forced to give up $2+2+3+4 = 11$ units and consequently are left with 9 units to distribute between the four opportunities. We can distribute 9 units between 4 opportunities in the total number of integer solutions to the linear equation $x_1 + x_2 + x_3 + x_4 = 9$ ways. The answer to that latter part is given by the solution to the stars and bars problem which in this case is

$$\binom{9+4-1}{4-1}.$$

For part (b), we have taken care of the case "investments must be made in more that 3 of the 4 opportunities" in part (a). Hence, we need only worry about the case "investments must be made in exactly 3 of the 4 opportunities." Since our unit total will change depending on which 3 opportunities we choose, we must consider these $\binom{4}{3} = 4$ cases one by one.

First, let us consider the case in which we choose the 2-3-4 opportunities. After the initial investment we have $20 - 4 - 3 - 2 = 11$ units left over giving us

$$\binom{11+3-1}{3-1}$$

different distributions of our remaining units among the 3 opportunities. We double this for the case in which we choose the other 2 unit opportunity.

Next we consider the 2-2-3 investment. If we go the 2-2-3 route, we have $20-3-2-2 = 13$ units left over after the initial investment and

$$\binom{13+3-1}{3-1}$$

distributions of our units.

Lastly, for the 2-2-4 route, we have $20 - 2 - 2 - 4 = 12$ units left over after the initial investment giving us

$$\binom{12+3-1}{3-1}$$

distributions of our units.

Therefore, there are

$$2\binom{11+3-1}{3-1} + \binom{13+3-1}{3-1} + \binom{12+3-1}{3-1} + \binom{9+4-1}{4-1}$$

ways of investing in at least 3 of the 4 opportunities. ∎

3.1.2 Theoretical exercises

PROBLEM 3.1.7 (Ross, §1, # 5). Determine the number of vectors $(x_1, \ldots, x_n)$ such that each x_k is either 0 or 1 and

$$\sum_{k=1}^{n} x_k \geq l.$$

SOLUTION. An implicit assumption must be that $n \geq l$ for otherwise $(1, \ldots, 1)$ would not satisfy $\sum_{k=1}^{n} x_k \geq l$. We henceforth assume $n \geq l$.

For a vector $(x_1, \ldots, x_n)$ to satisfy

$$\sum_{k=1}^{n} x_k \geq l$$

at least l of the x_k must be 1. One way to approach this is counting the different number of ways to achieve each number from l to n. There are n vectors of the form $(0, \ldots, 0, 1, 0, \ldots, 0)$ and we want to consider distinct sums of at least l of these. Therefore, we have

$$\sum_{k=l}^{n} \binom{n}{k}$$

different ways to achieve this. ∎

PROBLEM 3.1.8 (Ross, §1, # 6). How many vectors $(x_1, \ldots, x_k)$ are there for which each x_k is a positive integer such that $1 \leq x_j \leq n$ and $x_1 < x_2 < \cdots < x_k$?

SOLUTION. Ignoring degenerate cases such as would arise if $n < k$, let us analyze the situation presented in the statement of the problem for $n \geq k$.

There are

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

solutions ∎

PROBLEM 3.1.9 (Ross, §1, # 8). Prove that

$$\binom{n+m}{r} = \binom{n}{0}\binom{m}{r} + \binom{n}{1}\binom{m}{r-1} + \cdots + \binom{n}{r}\binom{m}{0}.$$

Solution. ■

Problem 3.1.10 (Ross, §1, # 9). Use Theoretical Exercise 8 to prove that

$$\binom{2}{n} = \sum_{k=0}^{n} \binom{n}{k}^2.$$

Solution. ■

Problem 3.1.11 (Ross, §1, # 12). Consider the following combinatorial identity:

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}.$$

(a) Present a combinatorial argument for this identity by considering the set of n people and determining, in two ways, the number of possible selections of a committee of any size and a chairperson for the committee.

Hint:

(i) How many possible elections are there of a committee of size k and its chairperson?

(ii) How many possible selections are there of a chairperson and the other committee members?

(b) Verify the following identity for $n = 1, 2, 3, 4, 5$:

$$\sum_{k=1}^{n} \binom{n}{k} k^2 = 2^{n-2}n(n+1).$$

For a combinatorial proof of the preceding, consider the set of n people and argue that both sides of the identity represent the number of different selections of a committee, its chairperson, and its secretary (possibly the same as the chairperson).

Hint:

(i) How many different selection result in the committee containing exactly k people?

(ii) How many different selections are there in which the chairperson and the secretary are the same? (Answer: $n2^{n-1}$.)

(iii) How many different selections result in a chairperson and the secretary being different?

(c) Now argue that

$$\sum_{k=1}^{n} \binom{n}{k} k^3 = 2^{n-3}n^2(n+3).$$

Solution. ■

PROBLEM 3.1.12 (Ross, §1, # 23). Determine the number of vectors $(x_1, \ldots, x_n)$ of n variables such that each x_k is a nonnegative integer and

$$\sum_{k=1}^{n} x_k \leq l.$$

SOLUTION. ■

3.1.3 Problems

PROBLEM 3.1.13 (Ross, §2, # 25). A pair of dice is rolled until a sum of either 5 or 7 appears. Find the probability that a 5 occurs first.
Hint: Let E_n denote the event that a 5 occurs on the n^{th} roll and no 5 or 7 occurs on the first $n-1$ rolls. Compute $P(E_n)$ and argue that $\sum_{n=1}^{\infty} P(E_n)$ is the desired probability.

SOLUTION. ■

PROBLEM 3.1.14 (Ross, §2, # 29). An urn contains n white and m black balls, where n and m are positive numbers.

(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).

SOLUTION. ■

PROBLEM 3.1.15 (Ross, §2, # 35). Seven balls are randomly withdrawn from an urn that contains 12 red, 16 blue, and 18 green balls. Find the probability that

(a) 3 red, 2 blue, and 2 green balls are withdrawn;

(b) at least 2 red balls are withdrawn;

(c) all withdrawn balls are the same color;

(d) either exactly 3 red balls or exactly 3 blue balls are withdrawn.

SOLUTION. ■

PROBLEM 3.1.16 (Ross, §2, # 44). Five people, designated as A, B, C, D, E, are arranged in linear order. Assuming that each possible order is equally likely, what is the probability that

(a) there is exactly one person between A and B?

(b) there are exactly two people between A and B?

(c) there are three people between A and B?

Solution. ■

Problem 3.1.17 (Ross, §2, # 49). A group of 6 men and 6 women is randomly divided into 2 groups of size 6 each. What is the probability that both groups will have the same number of men?

Solution. ■

3.1.4 Theoretical exercises

Problem 3.1.18 (Ross, §2, # 5). For any sequence of events $E_1, E_2, \ldots$, define a new sequence $F_1, F_2, \ldots$, of disjoint events (that is, events such that $F_j \cap F_k = \varnothing$ whenever $j \neq k$) such that for all $n \geq 1$,

$$\bigcup_{k=1}^{n} F_k = \bigcup_{k=1}^{n} E_k.$$

Solution. ■

Problem 3.1.19 (Ross, §2, # 14). Prove Proposition 4.4 by mathematical induction.

Solution. Proposition 4.4 is called the inclusion-exclusion identity (or inclusion-exclusion principle which the author prefers to use).

Proposition (Inclusion-exclusion identity)**.**

$$\begin{aligned} P(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{k=1}^{n} P(E_k) - \sum_{k_1<k_2} P(E_{k_1} \cap E_{k_2}) \\ + \cdots + (-1)^{r+1} \sum_{k_1<k_2<\cdots<k_r} P(E_{k_1} \cap \cdots \cap E_{k_r}) \\ + \cdots + (-1)^{n+1} P(E_1 \cap \cdots \cap E_n). \end{aligned}$$

The summation $\sum_{k_1<k_2<\cdots<k_r} P(E_{k_1} \cap \cdots \cap E_{k_r})$ *is taken over all the* $\binom{n}{r}$ *possible subsets of size* r *of the set* $\{1, \ldots, n\}$*.*

■

Problem 3.1.20 (Ross, §2, # 19). An urn contains n red and m blue balls. They are withdrawn one at a time until a total of r, $r \leq n$, red balls have been withdrawn. Find the probability that a total of k balls are withdrawn.
Hint: A total of k balls will be withdrawn if there are $r-1$ red balls in the first $k-1$ withdrawals and the k^{th} withdrawal is a red ball.

Solution. ■

3.2 Homework 2

3.2.1 Problems

Problem 3.2.1 (Ross, §3, # 5).

Solution. ■

Problem 3.2.2 (Ross, §3, # 6).

Solution. ■

Problem 3.2.3 (Ross, §3, # 7).

Solution. ■

Problem 3.2.4 (Ross, §3, # 15).

Solution. ■

Problem 3.2.5 (Ross, §3, # 17).

Solution. ■

Problem 3.2.6 (Ross, §3, # 18).

Solution. ■

Problem 3.2.7 (Ross, §3, # 22).

Solution. ■

Problem 3.2.8 (Ross, §3, # 25).

Solution. ■

Problem 3.2.9 (Ross, §3, # 27).

Solution. ■

Problem 3.2.10 (Ross, §3, # 32).

Solution. ∎

Problem 3.2.11 (Ross, §3, # 37).

Solution. ∎

Problem 3.2.12 (Ross, §3, # 46).

Solution. ∎

Problem 3.2.13 (Ross, §3, # 55).

Solution. ∎

Problem 3.2.14 (Ross, §3, # 64).

Solution. ∎

Problem 3.2.15 (Ross, §3, # 86).

Solution. ∎

3.2.2 Theoretical exercises

Problem 3.2.16 (Ross, §3, # 3).

Solution. ∎

Problem 3.2.17 (Ross, §3, # 7).

Solution. ∎

Problem 3.2.18 (Ross, §3, # 12).

Solution. ∎

Problem 3.2.19 (Ross, §3, # 21).

Solution. ∎

Problem 3.2.20 (Ross, §3, # 22).

Solution. ∎

3.3 Homework 3

3.3.1 Problems

Problem 3.3.1 (Ross, §4, # 4).

Solution. ■

Problem 3.3.2 (Ross, §4, # 21).

Solution. ■

Problem 3.3.3 (Ross, §4, # 22).

Solution. ■

Problem 3.3.4 (Ross, §4, # 36).

Solution. ■

Problem 3.3.5 (Ross, §4, # 37).

Solution. ■

Problem 3.3.6 (Ross, §4, # 42).

Solution. ■

Problem 3.3.7 (Ross, §4, # 43).

Solution. ■

3.3.2 Theoretical exercises

Problem 3.3.8 (Ross, §4, # 4).

Solution. ■

Problem 3.3.9 (Ross, §4, # 5).

Solution. ■

Problem 3.3.10 (Ross, §4, # 10).

Solution. ■

Problem 3.3.11 (Ross, §4, # 13).

Solution. ■

Problem 3.3.12 (Ross, §4, # 14).

Solution. ■

Problem 3.3.13 (Ross, §4, # 18).

Solution. ■

Problem 3.3.14 (Ross, §4, # 25).

Solution. ■

Problem 3.3.15 (Ross, §4, # 27).

Solution. ■

Problem 3.3.16 (Ross, §4, # 30).

Solution. ■

Problem 3.3.17 (Ross, §4, # 32).

Solution. ■

Problem 3.3.18 (Ross, §4, # 34).

Solution. ■

3.3.3 Self-test problems

PROBLEM 3.3.19. Let X be a positive-integer valued random varibale; i.e., X takes values in $\mathbb{Z}^+$. Suppose the distribution of X satisfies the following: for all m, n it holds that

$$P(X > n) = P(X > m + n).$$

(The above property is called the memoryless property.)

Show that X must be a geometric random variable; i.e., there exists a p $(0 < p < 1)$ such that for all n

$$P(X = n) = (1 - p)^{n-1} p.$$

(Combined with what is shown in class, you will have actually proved that a positive-integer valued random variable is geometric *if and only if* it satisfies the memoryless property.)

SOLUTION. ■

3.4 Homework 4

3.4.1 Problems

PROBLEM 3.4.1 (Ross, §5, # 4).

SOLUTION. ■

PROBLEM 3.4.2 (Ross, §5, # 13).

SOLUTION. ■

PROBLEM 3.4.3 (Ross, §5, # 31).

SOLUTION. ■

PROBLEM 3.4.4 (Ross, §5, # 38).

SOLUTION. ■

PROBLEM 3.4.5 (Ross, §5, # 39).

SOLUTION. ■

PROBLEM 3.4.6 (Ross, §5, # 41).

SOLUTION. ■

PROBLEM 3.4.7 (Ross, §5, # 43).

SOLUTION. ■

3.4.2 Theoretical exercises

PROBLEM 3.4.8 (Ross, §5, # 1).

SOLUTION. ■

PROBLEM 3.4.9 (Ross, §5, # 3).

SOLUTION. ■

Problem 3.4.10 (Ross, §5, # 5).

Solution. ■

Problem 3.4.11 (Ross, §5, # 8).

Solution. ■

Problem 3.4.12 (Ross, §5, # 19).

Solution. ■

Problem 3.4.13 (Ross, §5, # 29).

Solution. ■

Problem 3.4.14 (Ross, §5, # 25).

Solution. ■

Problem 3.4.15 (Ross, §5, # 27).

Solution. ■

3.4.3 Self test

Problem 3.4.16 (Ross, §5, # 16).

Solution. ■

Problem 3.4.17 (Ross, §5, # 18).

Solution. ■

Problem 3.4.18 (Ross, §5, # 22).

Solution. ■

3.5 Homework 5

3.5.1 Problems

PROBLEM 3.5.1 (Ross, §5, # 16).

SOLUTION. ■

PROBLEM 3.5.2 (Ross, §5, # 20).

SOLUTION. ■

PROBLEM 3.5.3 (Ross, §5, # 23).

SOLUTION. ■

PROBLEM 3.5.4 (Ross, §5, # 30).

SOLUTION. ■

3.5.2 Theoretical exercises

PROBLEM 3.5.5 (Ross, §5, # 11).

SOLUTION. ■

PROBLEM 3.5.6 (Ross, §5, # 31).

SOLUTION. ■

3.5.3 Self-test problems

PROBLEM 3.5.7. Let $f(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \mathrm{e}^{-x^2/(2\sigma_1^2)}$ and $g(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \mathrm{e}^{-x^2/(2\sigma_2^2)}$. Show that

$$\int_{-\infty}^{\infty} f(y)g(x-y)\,dy = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \mathrm{e}^{-x^2/2(\sigma_1^2+\sigma_2^2)}.$$

Hint: there are many methods to do this problem. One, outlined in class is to first combine the exponentials and then compute the square for the quadratic function in the exponential.

SOLUTION. ■

3.6 Homework 6

3.6.1 Problems

PROBLEM 3.6.1 (Ross, §6, # 1).

SOLUTION. ■

PROBLEM 3.6.2 (Ross, §6, # 8).

SOLUTION. ■

PROBLEM 3.6.3 (Ross, §6, # 9).

SOLUTION. ■

PROBLEM 3.6.4 (Ross, §6, # 17).

SOLUTION. ■

PROBLEM 3.6.5 (Ross, §6, # 19).

SOLUTION. ■

PROBLEM 3.6.6 (Ross, §6, # 23).

SOLUTION. ■

PROBLEM 3.6.7 (Ross, §6, # 27).

SOLUTION. ■

PROBLEM 3.6.8 (Ross, §6, # 40).

SOLUTION. ■

PROBLEM 3.6.9 (Ross, §6, # 41).

SOLUTION. ■

Problem 3.6.10 (Ross, §6, # 42).

Solution. ■

Problem 3.6.11 (Ross, §6, # 43).

Solution. ■

Problem 3.6.12 (Ross, §6, # 54).

Solution. ■

Problem 3.6.13 (Ross, §6, # 55).

Solution. ■

Problem 3.6.14 (Ross, §6, # 56).

Solution. ■

Problem 3.6.15 (Ross, §6, # 57).

Solution. ■

3.7 Homework 7

3.7.1 Problems

Problem 3.7.1 (Ross, §7, # 16).

Solution. ■

Problem 3.7.2 (Ross, §7, # 25).

Solution. ■

Problem 3.7.3 (Ross, §7, # 26).

Solution. ■

Problem 3.7.4 (Ross, §7, # 38).

Solution. ■

Problem 3.7.5 (Ross, §7, # 40).

Solution. ■

Problem 3.7.6 (Ross, §7, # 70).

Solution. ■

Problem 3.7.7 (Ross, §7, # 71).

Solution. ■

Problem 3.7.8 (Ross, §7, # 72).

Solution. ■

3.7.2 Theoretical exercises

PROBLEM 3.7.9 (Ross, §7, # 10).

Solution. ■

PROBLEM 3.7.10 (Ross, §7, # 17).

Solution. ■

PROBLEM 3.7.11 (Ross, §7, # 38).

Solution. ■

PROBLEM 3.7.12 (Ross, §7, # 41).

Solution. ■

3.7.3 Problems

PROBLEM 3.7.13 (Ross, §8, # 1).

Solution. ■

PROBLEM 3.7.14 (Ross, §8, # 2).

Solution. ■

PROBLEM 3.7.15 (Ross, §8, # 3).

Solution. ■

PROBLEM 3.7.16 (Ross, §8, # 4).

Solution. ■

PROBLEM 3.7.17 (Ross, §8, # 9).

Solution. ■

PROBLEM 3.7.18 (Ross, §8, # 13).

Solution. ■

3.7.4 Theoretical exercises

PROBLEM 3.7.19 (ROSS, §8, # 6).

SOLUTION. ∎

PROBLEM 3.7.20 (ROSS, §8, # 7).

SOLUTION. ∎

4 Midterms and Finals (Yip)

4.1 Midterm 1

The author of this document apologizes for any ungrammatical content in this document. We present the problems as they are given to us and only correct grammar and orthography when it is obvious what the author meant.

PROBLEM 4.1.1. Suppose n balls are distributed at random into r boxes in such a way that each ball chooses a box independently of each other. Let S be the number of *empty boxes*. Compute $E(S)$ and $\mathrm{Var}(S)$.
Hint: Consider the random variables X_k, for $k = 1, \ldots, r$, which equals 1 if the k^{th} box is empty and 0 otherwise. Relate S and the X_k.

SOLUTION. Write $S = \sum_{k=1}^{n} X_k$ where

$$X_k = \begin{cases} 1 & \text{if the } k^{\text{th}} \text{ box is empty,} \\ 0 & \text{otherwise.} \end{cases}$$

Then, by the definition of expected value, we have

$$\begin{aligned} E(S) &= E\Big(\sum\nolimits_{k=1}^{r} X_k\Big) \\ &= \sum_{k=1}^{n} E(X_k) \end{aligned}$$

which, by the tail-sum formula, becomes

$$= rE(X_1).$$

Hence, to compute $E(S)$ we need first compute $E(X_1)$, as we now do. Unraveling the definition of expected value for X_1 we have

$$E(X_1) = 0 \cdot P(X_1 = 0) + 1 \cdot P(X_1 = 1) = P(X_1 = 1)$$

where

$$P(X_1 = 1) = \frac{(r-1)^n}{r^n}.$$

Thus,

$$E(S) = \frac{r(r-1)^n}{r^n} = \frac{(r-1)^n}{r^{n-1}}.$$

As for the variance, the second moment can likewise be computed as

$$\begin{aligned}
E(S^2) &= E\left[\left(\sum\nolimits_{k=1}^{r} X_k\right)^2\right] \\
&= E\left(\sum\nolimits_{k=1}^{r} X_k^2\right) + E\left(\sum\nolimits_{j\neq k}^{r} X_j X_k\right) \\
&= \sum_{k=1}^{r} E(X_k^2) + \sum_{j\neq k} E(X_j X_k) \\
&= \sum_{k=1}^{r} E(X_k) + \sum_{j\neq k} E(X_j X_k) \\
&= \frac{(r-1)^n}{r^{n-1}} + \sum_{j\neq k} E(X_j X_k).
\end{aligned}$$

Now, the sum

$$\sum_{j\neq k} E(X_j X_k) = \sum_{j\neq k} P(X_j = 1, X_k = 1) = \sum_{j\neq k} \frac{(r-2)^n}{r^n} = \binom{r}{2} \frac{(r-2)^n}{r^n}.$$

Thus,

$$\mathrm{Var}(S) = \frac{(r-1)^n}{r^{n-1}} + \binom{r}{2} \frac{(r-2)^n}{r^n} - \frac{(r-1)^{2n}}{r^{2n-2}}. \qquad \blacksquare$$

PROBLEM 4.1.2. Suppose n balls are distributed in n boxes in such a way that each ball chooses a box independently of each other.

(a) What is the probability that box # 1 is empty?

(b) What is the probability that only box # 1 is empty?

(c) What is the probability that only one box is empty?

(d) Given that box # 1 is empty, what is the probability that only one box is empty?

(e) Given that only one box is empty, what is the probability that box # 1 is empty?

Hint: make use of the fact that the number of balls and boxes are the same.

SOLUTION. For part (a), the probability that box # 1 is empty can be quickly shown to be

$$\frac{(n-1)^n}{n^n}; \tag{4.1.1}$$

i.e., there are n^n ways to distribute the n balls among the n boxes; then there are $(n-1)^n$ ways to distribute the n balls among the $n-1$ remaining boxes.

For part (b), the probability that only box # 1 is empty is

$$\frac{\binom{n}{2}(n-1)!}{n^n}; \tag{4.1.2}$$

i.e., each box must contain at least 1 ball and there are $\binom{n}{2}(n-1)!$ ways to distribute 2 of the n balls among the remaining $n-1$ boxes.

For part (c), using (4.1.2), the probability that exactly one box is empty is

$$\frac{n\binom{n}{2}(n-1)!}{n^n};$$

i.e., there are $\binom{n}{1} = n$ ways to choose a box to be empty once this is done, the probability that said box is empty is the same as the probability that box # 1 is empty; and that probability we found in part (b).

For part (d), by the definition of conditional probability together with parts (4.1.1) and (4.1.2) we have

$$\frac{\binom{n}{2}(n-1)!/n^n}{(n-1)^n/n^n} = \frac{\binom{n}{2}(n-1)!}{(n-1)^n}.$$

For part (e), the probability is clearly

$$\frac{1}{n}$$

since at most one box is empty and the event that box # 1 is empty is one sample point in our conditioned probability space. ∎

Problem 4.1.3. McDonald's newest promotion is putting a toy inside every one of its hamburgers. Suppose there are N distinct types of toys and each of them is equally likely to be put inside any of the hamburges. What is the expected value and variance of the number of hamburgers you need to order (or eat) before you have a complete set of the N toys.
Hint: consider the number of hamburgers you need to order (or eat) in between getting one and two distinct types of toys, two and three distinct types of toys, and so forth.

Solution. Setting the utter wackiness of this problem aside for a moment, let X_k denote the number of burgers we need to purchase in order to get the k^{th} toy, for $1 \leq k \leq N$. Before preceding, note that each X_k, for $k > 1$, is a Geo($\frac{N-k+1}{N}$) random variable; i.e., in our j^{th} try we succeed with in getting a new toy with probability $(\frac{N-k+1}{N})^j$. Then the total number of burgers we must purchase is

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

Thus,

$$\begin{aligned} E(X) &= \sum_{k=1}^{n} E(X_k) \\ &= 1 + \frac{N}{N-1} + \cdots + \frac{N}{1} \\ &= N\left(\frac{1}{N} + \frac{1}{N-1} + \cdots + 1\right). \end{aligned}$$

For the variance, we have

$$
\begin{aligned}
\operatorname{Var}(X) &= \sum_{k=1}^{n} \operatorname{Var}(X_k) \\
&= 0 + \frac{1/N}{((N-1)/N)^2} + \cdots + \frac{(N-1)/N}{(1/N)^2} \\
&= N\left(\frac{1}{(N-1)^2} + \frac{2}{(N-2)^2} + \cdots + N - 1\right). \qquad \blacksquare
\end{aligned}
$$

Problem 4.1.4. Let X and Y be two independent geometric random variables with parameter p.

(a) Find the probability distribution function of $\min\{X, Y\}$.

(b) Find the probability distribution function of $\max\{X, Y\}$.

(c) Find the probability distribution function of $X + Y$.

(d) Find $P(X = j \mid X + Y = k)$ for $j = 1, \dots, k-1$.

Solution. For part (a), let $Z = \min\{X, Y\}$. Then using the axioms of probability, we can manipulate the expression for the PMF of Z into the following

$$
\begin{aligned}
&P(Z = k) \\
&\qquad = P(X = Y = k) + P(X = k, Y > k) + P(X > k, Y = k) \qquad (4.1.3) \\
&\qquad = P(X = k)P(Y = k) + P(X = k)P(Y > k) + P(X > k)P(Y = k).
\end{aligned}
$$

Now let us find a closed form for the expression above. To make the analysis more digestible, we consider each term in the sum (4.1.3) individually.

First, we have

$$
\begin{aligned}
P(X = k)P(Y = k) &= (q^{k-1}p)(q^{k-1}p) \\
&= q^{2k-2}p^2
\end{aligned}
$$

Next,

$$
\begin{aligned}
P(X = k)P(Y > k) &= q^{k-1}p \sum_{j \geq k}^{\infty} q^{j-1}p \\
&= \sum_{j > k}^{\infty} (q^{2k-1}p^2) q^{j-k-1} \\
&= q^{2k-1}p.
\end{aligned}
$$

Lastly, because X and Y are identically distributed, by symmetry

$$
P(X > k)P(Y = k) = q^{2k-1}p.
$$

Therefore, the PMF of Z is given by the expression

$$P(Z = k) = q^{2k-2}p^2 + 2q^{2k-1}p.$$

For part (b), let $W = \max\{X, Y\}$. Then we can expand the PMF of W to get a simpler expression as follows

$$\begin{aligned} P(W = k) &= P(X = Y = k) + P(X = k, Y < k) + P(X < k, Y = k) \\ &= P(X = k)P(Y = k) + 2P(X = k)P(Y < k). \end{aligned} \tag{4.1.4}$$

One term of the expression (4.1.4) we already know from part (a); namely,

$$P(X = k)P(Y = k) = q^{2k-2}p^2.$$

As for the other term, we have

$$\begin{aligned} P(X = k)P(Y < k) &= q^{k-1}p\sum_{j=1}^{k} q^{j-1}p \\ &= q^{k-1}\frac{1-q^k}{p}p^2 \\ &= q^{k-1}(1-q^k)p. \end{aligned}$$

Thus, the PMF of W is given by the expression

$$P(W = k) = q^{2k-2}p^2 + 2q^{k-1}(1-q^k)p.$$

For part (c), since X and Y are independent the PMF of $X + Y$ is given by the convolution

$$P(X + Y = k) = \sum_{j=1}^{k-1} P(X = k - j)P(Y = j).$$

Let us find a closed form for this expression,

$$\begin{aligned} P(X + Y = k) &= \sum_{j=1}^{k-1} P(X = k - j)P(Y = j) \\ &= \sum_{j=1}^{k-1}(q^{(k-j)-1}p)(q^{j-1}p) \\ &= \sum_{j=1}^{k-1} q^{k-2}p^2 \\ &= (k-1)q^{k-2}p^2. \end{aligned}$$

For part (d), by the definition of conditional probability, we have

$$P(X = j \mid X + Y = k) = \frac{P(X = j, X + Y = k)}{P(X + Y = k)} = \frac{P(X = j)P(Y = k - j)}{P(X + Y = k)}. \tag{4.1.5}$$

By our answer to part (c) together with the PMFs of X and Y, (4.1.5) becomes

$$\begin{aligned} P(X = j \mid X + Y = k) &= \frac{(q^{j-1}p)(q^{k-j-1}p)}{(k-1)q^{k-2}p^2} \\ &= \frac{1}{k-1}. \end{aligned}$$

■

PROBLEM 4.1.5. Suppose the events E, F, G are independent, in other words

$$\begin{aligned} P(E \cap F) &= P(E)P(F), \\ P(F \cap G) &= P(F)P(G), \\ P(G \cap E) &= P(G)P(E), \\ P(E \cap F \cap G) &= P(E)P(F)P(G). \end{aligned}$$

Using he above definition, show that the following events are independent:

(a) E and F^c;

(b) E and $F \cap G^c$;

(c) E and $F^c \cap G^c$;

(d) E and $F \cup G$;

(e) E and $F \cup G^c$.

SOLUTION. For part (a), we have

$$\begin{aligned} P(E \cap F^c) &= P(E) - P(E \cap F) \\ &= P(E) - P(E)P(F) \\ &= P(E)(1 - P(F)) \\ &= P(E)P(F^c). \end{aligned}$$

For part (b), we have

$$\begin{aligned} P(E \cap (F \cap G^c)) &= P((E \cap F) \cap (E \cap G^c)) \\ &= P(E \cap F) - P(E \cap G) \\ &= P(E)P(F) - P(E)P(G) \\ &= P(E)(P(F) - P(G)) \\ &= P(E)P(F \cap G^c). \end{aligned}$$

For part (c), we have

$$\begin{aligned}
P(E \cap F^c \cap G^c) &= P((E \cap F^c) \cap (E \cap G^c)) \\
&= P(E \cap F^c) - P(E \cap G) \\
&= P(E)P(F^c) - P(E)P(G) \\
&= P(E)(P(F^c) - P(G)) \\
&= P(E)(P(F^c \cap G^c)).
\end{aligned}$$

For part (d), we have

$$\begin{aligned}
P(E \cap F \cup G) &= P((E \cap F) \cup (E \cap G)) \\
&= P(E \cap F) + P(E \cap G) - P(E \cap F \cap G) \\
&= P(E)P(F) + P(E)P(G) - P(E)P(F)P(G) \\
&= P(E)(P(F) + P(G) - P(F)P(G)) \\
&= P(E)(P(F) + P(G) - P(F \cap G)) \\
&= P(E)P(F \cup G).
\end{aligned}$$

For part (e), we have

$$\begin{aligned}
P(E \cap F \cup G^c) &= P((E \cap F) \cup (E \cap G^c)) \\
&= P(E \cap F) + P(E \cap G^c) - P(E \cap F \cap G^c) \\
&= P(E)P(F) + P(E)P(G^c) - P(E)P(F)P(G^c) \\
&= P(E)(P(F) + P(G^c) - P(F)P(G^c)) \\
&= P(E)(P(F) + P(G^c) - P(F \cap G^c)) \\
&= P(E)P(F \cup G^c).
\end{aligned}$$

■

4.2 Midterm 2

This is actually the past final.

PROBLEM 4.2.1. Consider the infinitely many independent, identical experiments of throwing a pair of dice and the outcome of each experiment is the sum of the two numbers. Let $N \geq 1$ be the number of experiments such that the number 5 or 7 first appears as the outcome. Let further X be that number at the N^{th} experiment; i.e., X can be either a 5 or a 7.

(a) Find $P(N = n)$ for $n \geq 1$.

(b) Find $P(X = 5)$.

(c) Prove or disprove that N and X are independent.

Solution. For part (a): It is clear that $N \sim \text{Geo}(p)$ where p is the probability of getting a 5 or a 7 in a trial of the experiment in question. Let us therefore compute p. Assuming each event is equally likely, there are 4 ways to make a 5 (rolling a 4-1 or 2-3 and vice-versa) and 6 ways to make a 7 (rolling a 6-1, 5-2, or 4-3 and vice-versa). Therefore, $p = \frac{4}{36} + \frac{6}{36} = \frac{5}{18}$ and consequently the PMF of N is given by

$$P(N = n) = \left(1 - \tfrac{5}{18}\right)^{n-1} \tfrac{5}{18} = \left(\tfrac{13}{18}\right)^{n-1} \tfrac{5}{18}.$$

For part (b): Since X takes the values 5 and 7 conditioned upon the result of a trial of a pair of dice by the definition of conditional probability we have

$$P(X = 5) = \frac{P(\text{the experiment returns a 5})}{P(\text{the experiment returns either a 5 or a 7})} = \frac{4/36}{10/36} = \frac{2}{5}.$$

For part (c): To prove that N and X are independent, we must show that

$$P(N = n, X = x) = P(N = n)P(X = x).$$

Since $X = \text{Ber}(\frac{2}{5})$, we only have a couple of cases to consider. For $X = 5$ consider

$$P(N = n, X = 5) = \left(1 - \tfrac{2}{18}\right)^{n-1} \tfrac{2}{18}$$

■

PROBLEM 4.2.2. Consider a city in which the male and female drivers occupy α and $1 - \alpha$ fractions of the whole city driver population. In any given year, a male and female driver will have an accident with probability p_M and p_F. Assume that the behavior of each driver is independent from year to year.

Now a driver is randomly chosen. Let A_k be the event that this driver will have an accident in the k^{th} year. Let M be the event that the randomly chosen driver is male.

(a) Suppose $p_M > p_F$. Show that $P(M \mid A_k) > p(M)$.

(b) Suppose $p_M \neq p_F$. Show that $P(A_2 \mid A_1) > p(A_1)$.

Solution. ∎

Problem 4.2.3. Let $X_1, \dots, X_n$ be a collection of IID exponential random variables with parameter λ. Let

$$\begin{aligned} Y_1 &= X_1, \\ Y_2 &= X_1 + X_2, \\ &\vdots \\ Y_n &= X_1 + \dots + X_n. \end{aligned}$$

Find the joint PDF $p(y_1, \dots, y_n)$ of $Y_1, \dots, Y_n$.

Solution. ∎

Problem 4.2.4. The PDF $p(x)$ of the gamma distribution with parameter ($\alpha > 0$, $\lambda > 0$) is given by

$$p(x) = \begin{cases} \frac{\lambda}{\Gamma(\alpha)} \mathrm{e}^{-\lambda x} (\lambda x)^{\alpha - 1} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let X and Y be independent gamma distributed random variables with parameters (α, λ) and (β, λ). Show analitically that $X + Y$ has a gamma distribution with parameter $(\alpha + \beta, \lambda)$.

Show how as a byproduct that the above conclusion leads to the following integration identity for $\alpha, \beta > 0$

$$\int_0^1 x^{\alpha - 1} (1 - x)^{\beta - 1} \, dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Solution. ∎

Problem 4.2.5. Let $X_1, \dots, X_n$ be a collection of IID random variables with expectations and variances equal to μ and σ^2. Define the *sample mean* $\bar{X}$ and *sample variance* S^2 as

$$\bar{X} = \tfrac{1}{n}(X_1 + \dots + X_n), \quad S^2 = \frac{1}{n} \sum_{k=1}^{n} (X_k - \bar{X})^2.$$

Compute $\mathrm{Var}(\bar{X})$ and $E(S^2)$.

Solution. ∎

PROBLEM 4.2.6 (Estimation of the length of an interval). Let $l > 0$ be some unknown but fixed length. Let $X_1, X_2, \ldots$, be a sequence of IID random variables uniformly distributed on $[0, l]$. The goal is to use the X_k to estimate l.

(a) Let $A_n = \frac{2}{n}(X_1 + \cdots + X_n)$. Show that A_n is an unbiased estimation in the sense that $E(A_n) = l$.

(b) Let $B_n = \gamma_n \max\{X_1, \ldots, X_n\}$ where γ_n is some number. Find the correct value of γ_n such that B_n is also an unbiased estimator.

(c) Find $\mathrm{Var}(A_n)$ and $\mathrm{Var}(B_n)$.

(d) Which

SOLUTION. ■

5 Midterms, Exams, and Qualifying Exams

These are das Gupta's qualifying exams.

5.1 Qualifying Exams, August '99

Problem 5.1.1. The number of fish that Anirban catches on any given day has a Poisson distribution with mean 20. Due to the legendary softness of his heart, he sets free, on average, 3 out of the 4 fish he catches. Find the mean and the variance of the number of fish Anirban takes home on a given day.

Solution. Let X denote the number of fish caught by Anirban on any given day and let Y denote the number of fish released by Anirban. Since Anirban releases on average three-fourths of the fish he catches, the number of fish he keeps is

$$K := X - Y = X - \tfrac{3}{4}Y = \tfrac{1}{4}X.$$

Therefore,

$$E(K) = \tfrac{1}{4}E(X) = \frac{20}{4} = 5$$

and

$$\operatorname{Var}(K) = \left(\tfrac{1}{4}\right)^2 \operatorname{Var}(X) = \frac{20}{16}.$$

■

Problem 5.1.2. A fair die is rolled and at the same time a fair coin is tossed. This is done repeatedly. Find the probability that head occurs (strictly) before six occurs.

Solution. Let X denote the number of tosses until a head comes up and Y denote the number of rolls until we roll a six. Both of these random variables have geometric PMFs with parameters $\frac{1}{2}$ and $\frac{1}{6}$, respectively. Then we need to find $P(X < Y)$. Since X and Y are independent this value is given by the sum

$$
\begin{aligned}
P(X < Y) &= P(0 < Y - X) \\
&= \sum_{k=1}^{\infty} \sum_{l=k+1}^{\infty} P(X = k)P(Y = l) \\
&= \sum_{k=1}^{\infty} \sum_{l=k+1}^{\infty} \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{k-1} \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{l-1} \\
&= \sum_{k=1}^{\infty} \sum_{l=k+1}^{\infty} \\
&= \left(\frac{1}{12}\right)\left[\sum_{k=0}^{\infty}\left(\frac{1}{12}\right)^{k-1}\right]\left[\sum_{l=0}^{\infty}\left(\frac{5}{6}\right)^{l}\right] \\
&= \left(\frac{1}{12}\right)\left(\frac{1}{1-\frac{1}{12}}\right)\left(\frac{1}{1-\frac{5}{6}}\right) \\
&= \frac{6}{11}.
\end{aligned}
$$

■

PROBLEM 5.1.3. X, Y are independent random variables with a common density $f(x) = \frac{e^{-|x|}}{2}$, $x \in (-\infty, \infty)$. Find the density function of $X + Y$.

SOLUTION. Suppose X and Y are both double-exponential random variables both having identical PDFs $f_X(x) = f_Y(x) = \frac{e^{-|x|}}{2}$. Then, since X and Y are independent, we have

$$
\begin{aligned}
P(X + Y \leq x) &= \int_{-\infty}^{\infty} f_X(x) f_Y(x-y)\, dy \\
&= \frac{e^{-|x|}}{4} \int_{-\infty}^{\infty} e^{-|x-y|}\, dy \\
&= \frac{e^{-|x|}}{4} \left[\int_{-\infty}^{x} e^{x-y}\, dy + \int_{x}^{\infty} e^{y-x}\, dy\right] \\
&\frac{e^{-|x|}}{4}[1+1] \\
&= \frac{e^{-|x|}}{2}.
\end{aligned}
$$

■

PROBLEM 5.1.4. Let X_n denote the distance between two points chosen independently at random from the unit cube in $\mathbb{R}^n$. Evaluate

$$\lim_{n\to\infty} \frac{E(X_n)}{\sqrt{n}}.$$

SOLUTION. The points, call them Y_n and Z_n are uniformly distributed on $[0, 1]^n$. First, note that if $a_n \to a$, $\sqrt{a_n} \to \sqrt{a}$. Let find a bound on the size of the expected value, by the

Cauchy–Schwartz inequality

$$\begin{aligned}
E(X_n^2) &= E\big((Y_{1,n} - Z_{1,n})^2 + \cdots + (Y_{n,n} - Z_{n,n})^2\big) \\
&= E\big(Y_{1,n}^2 - 2Y_{1,n}Z_{1,n} + Z_{1,n}^2 + \cdots + Y_{n,n}^2 - 2Y_{n,n}Z_{n,n} + Z_{n,n}^2\big) \\
&= E(Y_{1,n^2}) - 2E(Y_{1,n})E(Z_{1,n}) + E(Z_{1,n}^2) \\
&\quad + \cdots + E(Y_{n,n}^2) - 2E(Y_{n,n,})E(Z_{n,n}) + E(Z_{n,n}^2)
\end{aligned}$$

since Y_n and Z_n are independent, moreover, since each coordinate is uniformly distributed on $[0, 1]$, the equation above further reduces to

$$\begin{aligned}
E(X_n^2) &= 2nE(Y_{1,n}^2) - 2nE(Y_{1,n})E(Z_{1,n}) \\
&= 2n \operatorname{Var}(Y_{1,n}) \\
&= \frac{n}{6}.
\end{aligned}$$

By the strong law of large numbers

$$\frac{X_n^2}{n} \longrightarrow \frac{n/6}{n} = \frac{1}{6};$$

i.e., $X_n/n \to \frac{1}{\sqrt{6}}$. ■

Problem 5.1.5. Let X be distributed as $U[0, 1]$. What is the probability that the digit 5 does not occur in the decimal expansion of X?

Solution. ■

5.2 Qualifying Exam, January '06

PROBLEM 5.2.1. The birthdays of 5 people are known to fall in exactly 3 calendar months. What is the probability that exactly two of the 5 were born in January?

SOLUTION. ∎

PROBLEM 5.2.2. Coupons are drawn, independently, with replacement, one at a time, from a set of 10 coupons. Find, explicitly, the expected number of draws

(a) until the first draw coupon is drawn again;
(b) until a duplicate occurs.

SOLUTION. ∎

PROBLEM 5.2.3. Let N be a positive integer. Choose an integer at random from $\{1, \dots, N\}$. Let E be the event that your chosen random number is divisible by 3, and divisible by at least one of 4 and 6, but not divisible by 5. Find, explicitly, $\lim_{N\to\infty} P(E)$.

SOLUTION. ∎

PROBLEM 5.2.4. Anirban is driving his Dodge on a highway with 4 lanes each way. He is wired to change lanes every minute on the minute. He changes with equal probability to either adjacent lane if there are two adjacent lanes, and the successive changes are mutually independent. Find, explicitly, the probability that after 4 minutes, Anirban is back to the lane he started from

(a) if he started at an outside lane;
(b) if he started at an inside lane.

SOLUTION. ∎

PROBLEM 5.2.5. Burgess is going to Moose Pass, Alaska. He is driving his Dodge. He puts his car on cruise control at 70 mph. Gas stations are located every 30 miles, starting from his home. His car runs out of gas at a time distributed as an exponential with mean 4 hours. When that happens, he gets out, takes his bik out of his trunk, and bikes to the next gas station say M, at 10 mph. Let the time elapsed between when Burgess starts his trip and when he arrives at the gas station M be T. Find $E(T)$.

SOLUTION. ∎

PROBLEM 5.2.6. A fair coin is tossed n times. Suppose X heads are obtained. Given $X = x$, let Y be generated according to the Poisson distribution with mean x. Find the unconditional variance of Y, and then find the limit of the probability $P\big(|Y - n/2| > n^{3/4}\big)$, as $n \to \infty$.

Solution. ∎

Problem 5.2.7. Anirban plays a game repeatedly. On each play he wins an amount uniformly distributed in $(0,1)$ dollars, and then he tips the lady in charge of the game the square of the amount he has won. Then he plays again, tips again, and so on. Approximately calculate the probability that if he plays and tips six hundred times, his total winnings minus his total tips will exceed \$105.

Solution. ∎

Problem 5.2.8. Anirban's dog got mad at him and broke his walking cane, first uniformly into two peices, and then the long piece again uniformly into two pieces. Find the probability that Anirban can make a triangle out of the three pieces of his cane.

Solution. ∎

Problem 5.2.9. Suppose X, Y, Z are identically independently distributed Exp(1) random variables. Find the joint density of (X, XY, XYZ).

Solution. ∎

Problem 5.2.10. Let X be the number of kings and Y the number of hearts in a Bridge hand. Find the correlation between X and Y.

Solution. ∎

5.3 Qualifying Exam, August '14

PROBLEM 5.3.1.

(a) 3 balls are distributed one by one and at random in 3 boxes. What is the probability that exactly one box remains empty?
(b) n balls are distributed one by one and at random in n boxes. Find the probability that exactly one box remains empty.
(c) n balls are distributed one by one and at random in n boxes Find the probability that exactly two boxes remain empty.

SOLUTION. ■

PROBLEM 5.3.2. n players each roll a fair die. For any pair of players i, j, $i < j$, who roll the same number, the group is awarded one point.

(a) Find the mean of the total points of the group.
(b) Find the variance of the total points of the group.

SOLUTION. ■

PROBLEM 5.3.3. Suppose $X_1, X_2, \ldots$, is an infinite sequence of independently identically distributed Uniform$[0, 1]$ random variables. Find the limit

$$\lim_{n\to\infty} P\left[\frac{\left(\prod_{i=1}^n X_i\right)^{1/n}}{\left(\sum_{i=1}^n X_i\right)/n} > \frac{3}{4}\right].$$

SOLUTION. ■

PROBLEM 5.3.4. Suppose X is an exponential random variable with density $\mathrm{e}^{-x/\sigma_1}/\sigma_1$ and Y is another exponential random variable with density $\mathrm{e}^{-y/\sigma_2}/\sigma_2$, and that X, Y are independent.

(a) Find the CDF of $X/(X+Y)$.
(b) In the case $\sigma_1 = 2$, $\sigma_2 = 1$, find the mean of $X/(X+Y)$.

SOLUTION. ■

PROBLEM 5.3.5. Ten independently picked Uniform$[0, 100]$ numbers are each rounded to the nearest integer. Use the central limit theorem to approximate the probability that the sum of the ten rounded numbers equals the rounded value of the sum of the ten original numbers.

SOLUTION. ■

PROBLEM 5.3.6. Suppose for some given $m \geq 2$, we choose m independently identically distributed Uniform$[0,1]$ random variables $X_1, \ldots, X_m$. Let $X_{\min}$ denote their minimum and $X_{\max}$ denote their maximum. Now continue sampling $X_{m+1}, \ldots,$ from the Uniform$[0,1]$ density. Let N be the first index k such that X_{m+k} falls outside the interval $[X_{\min}, X_{\max}]$.

(a) Find a formula for $P(N > n)$ for a general n.
(b) Hence, explicitly find $E(N)$.

SOLUTION. ■

PROBLEM 5.3.7. A $G_{n,p}$ graph on n vartices is obtained by adding each of the $\binom{n}{2}$ possible edges into the graph mutually independently with probability p. If vertex subsets A, B both have k vertices, and each vertex A shares an edge with each vertex in B, but there are no edges among the vertices within A or within B, then A, B generate a complete bipartate subgraph of order k denoted as $K_{k,k}$.

(a) For a given n and p, find an expression for the expected number of complete bipartate subgraphs $K_{3,3}$ of order $k = 3$ in a $G_{n,p}$ graph.
(b) Let p_n denote the value of p for which the expected value in part (a) equals one. Identify constants α, β such that $\lim_{n\to\infty} n^\alpha p_n = \beta$.

SOLUTION. ■

References

[1] A. DasGupta. *Fundamentals of Probability: A First Course*. Springer Texts in Statistics. Springer New York, 2010.

[2] W. Feller. *An introduction to probability theory and its applications*. Number v. 1 in Wiley mathematical statistics series. Wiley, 1950.

[3] S.M. Ross. *A First Course in Probability*. Pearson Education, 2014.