Project C-B Project

Charlie Windolf

Contents

Chapter 1

Probability Theory: Exercises

1. *For each of the following experiments, describe the sample space.*

   a. *Toss a coin four times.*

      $\{H, T\}^4$.

   b. *Count the number of insect-damaged leaves on a plant.*

      $\{0, \dots, N\}$, where N is the number of leaves.

   c. *Measure the lifetime (in hours) of a particular brand of light bulb.*

      Assuming we only measure discrete hours, $\mathbb{Z}_{\geq 0}$.

   d. *Record the weights of 10-day-old rats.*

      $\mathbb{R}_{\geq 0}$?

   e. *Observe the proportion of defectives in a shipment of electronic components.*

      $[0, 1]$.

2. *Verify the following identities.*

   a. $A \setminus B = A \setminus (A \cap B) = A \cap B^C$.

      These kinds of basic set theory questions are always hilarious, because the expected proof always assumes the logical version of the same claim, and nobody knows whether it would even help to learn how to prove it from "first principles".

      Anyway, by definition,

$$\begin{aligned} A \setminus B &= \{x : x \in A \text{ and } x \notin B\} \\ &= A \cap B^C. \end{aligned}$$

Also,

$$A \cap B^C = A \cap (A^C \cup B^C) = A \cap (A \cap B)^C = A \setminus (A \cap B).$$

b. $B = (B \cap A) \cup B(\cap A^C)$.

By distributive property,

$$B = B \cap S = B \cap (A \cup A^C) = (B \cap A) \cup B(\cap A^C).$$

c. $B \setminus A = B \cap A^C$. Yikes, this is shown in (a.)?

d. $A \cup B = A \cup (B \cap A^C)$.

Indeed,

$$\begin{aligned} A \cup B &= \{x \in S : x \in A \text{ or } x \in B\} \\ &= \{x \in S : x \in A \text{ or } (x \notin A \text{ and } x \in B)\} \\ &= A \cup (B \cap A^C). \end{aligned}$$

3. *Finish the proof of Theorem 1.1.4. For any events A, B, C defined on a sample space S, show that: (a) set union and intersection are commutative, (b) associative, and (c) obey De Morgan's laws.*

   These follow directly from commutativity, associativity, and the De Morgan's laws of logical and and or, which can be shown by looking at truth tables.

4. *For events A and B, find formulas for the following events in terms of $P(A)$, $P(B)$, and $P(A \cap B)$.*

   a. *either A or B or both:* This is inclusion-exclusion:

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

   b. *either A or B but not both:*

   $$P(A \cup B) = P(A \cup B) - P(A \cap B) = P(A) + P(B) - 2P(A \cap B).$$

   c. *at least one of A or B:* This is (a.), no?

   d. *at most one of A or B:* $P(S \setminus (A \cup B)) + P(A \cup B) - P(A \cap B)$.

5. (*Paraphrasing, and excuse the bad biology here.*) *One third of twins are identical, the rest fraternal. Identical twins are same sex, male and female equally likely. For fraternal twins, the pairs are equally likely, so that MM and FF occur w.p. 1/4 and MF w.p. 1/2. Finally, 1 in 90 births yields twins. Let*

$$\begin{aligned} A &= \{\text{birth results in FF}\} \\ B &= \{\text{birth results in identical}\}. \\ C &= \{\text{birth results in twins}\}. \end{aligned}$$

a. *State, in words, the event $A \cap B \cap C$:* A U.S. birth results in identical twin females.

b. *Find $P(A \cap B \cap C)$:*

   Well, since $B \subset C$, and by the chain rule,

$$\begin{aligned} P(A \cap B \cap C) &= P(A \mid B \cap C)P(B \mid C)P(C) \\ &= P(A \mid B)P(B \mid C)P(C) \\ &= \frac{1}{4} frac13 \frac{1}{90} \\ &= \frac{1}{1080}. \end{aligned}$$

6. *Two pennies, one with $P(H) = u$ and the other with $P(H) = w$ are tossed independently. Let p_i be the probability of the event that i heads occur. Can u and w be chosen such that $p_0 = p_1 = p_2$.*

   Well, we have:

$$\begin{aligned} p_0 &= (1-u) + (1-w) - (1-u)(1-w) = 1 - uw, \\ p_1 &= u(1-w) + w(1-u) = u + w, \\ p_2 &= uw. \end{aligned}$$

   Then we need

$$1 - uw = uw \implies uw = \frac{1}{2} = u + w.$$

   Substituting for $u = \frac{1}{2} - w$, we have

$$uw = w(\tfrac{1}{2} - w) = \frac{w}{2} - w^2 = \frac{1}{2},$$

   which has no real solutions.

7. This question recalls an example with a dartboard that's hit w.p. 1, and where each of its five rings is hit w.p. proportional to its area, so that if we write $P(i \text{ points scored}) = p_i$, then

$$p_i = \frac{(6-i)^2 - (5-i)^2}{5^2}.$$

   Now, we imagine that the dartboard is mounted on a wall that's hit w.p. 1, with area A (seems like $A \geq 25\pi$).

a. *Find the new point probabilities, call them p_i'.* Well,

$$p_i' = \frac{\pi(6-i)^2 - \pi(5-i)^2}{A}$$

   for $1 \leq i \leq 5$, with $p_0' = 1 - 25\pi/A$.

b. *Show that* $P(i$ points on new board | board hit$) = p_i$. Indeed,

$$\begin{aligned} P(i \text{ points on new board} \mid \text{board hit}) &= \frac{p_i'}{P(\text{board hit})} \\ &= p_i' \frac{A}{25\pi} \\ &= p_i. \end{aligned}$$

11. Let S be a sample space.

    a. Show that $\{\varnothing, S\}$ is a σ-algebra. It remains to check that this is closed under countable unions, but it's even preserved under those.

    b. Show that 2^S is a σ-algebra. Well, $\varnothing \subset S$, and complements or even arbitrary unions of subsets of S cannot escape S.

    c. Let $\mathcal{A}, \mathcal{B}$ be σ-algebras over S. Then they both contain $\varnothing$ and S. Also, if A is in their intersection, then A^C must have also been in both of them. Same goes for (even arbitrary) unions.

12. *Together, finite additivity and continuity imply countable additivity.*

    a. *Claim:* Countable additivity implies finite additivity. Indeed, just take $A_{n+1}, A_{n+2}, \ldots$ to be $\varnothing$.

    b. *Claim:* Finite additivity and continuity imply countable additivity.

    Let $A_1, A_2, \ldots$ be disjoint elements of a σ-algebra $\mathcal{F}$, let $A = \bigcup_i A_i$, and let $B_n = \bigcup_{i=n}^{\infty} A_i$. By disjointness, $\bigcap_n B_n = \varnothing$, so that by continuity $P(B_n) \to 0$.

    But by finite additivity,

    $$P(A) = \sum_{i=1}^{n-1} P(A_i) + P(B_n)$$

    for all n. Since $P(B_n) \to 0$,

    $$\lim_{n\to\infty} \sum_{i=1}^{n-1} P(A_i) = P(A).$$

    (The odd thing to me is that we are allowed to assume $A \in \mathcal{F}$: I would be more convinced of the inevitability of countable additivity if we had shown first that the finite union property for σ-algebras together with some continuity property, say being closed under countable intersections for sequences of nested sets. However, that would probably be strong enough to prove the countable union property due to a π-λ scenario.)

13. *If* $P(A) = \frac{1}{3}$ *and* $P(B^C) = \frac{1}{4}$, *can* A *and* B *be disjoint?*

Note $P(B) = \frac{3}{4}$. But by Bonferroni,

$$P(A \cap B) = P(A) + P(B) - 1 = \frac{1}{12}.$$

Since $P(\varnothing) = 0$, $A \cap B \neq \varnothing$.

14. *Let S be a finite set of n elements. Then $|2^S| = 2^n$.*

    Indeed, we can establish a bijection C between 2^S and $\{0, \dots, 2^{n-1}\}$ as follows. Assume w.l.o.g. $S = \{1, \dots, n\}$, and for any $A \subset S$ take $C(A)$ to be the number whose ith binary bit is $\mathbb{1}_{i \in A}$.

15. Finishing the proof of Thm. 1.2.14 just takes a simple induction: consider the first i tasks as one task, and use the $k = 2$ case to add the $i + 1$th task.

18. *Place n balls at random in n cells. What is the probability that exactly one cell remains empty?*

    There are n^n total arrangements of balls, $(n-1)^n$ with the first cell empty, and $n(n-1)^n$ with at least one cell empty. But we need to ensure that the rest of the cells are not empty.

    That implies that $n-2$ cells have 1 ball, 1 cell has 2 balls, and the remaining has nothing. Let the first cell be empty, which happens in 1 way. There are $\binom{n}{2}$ ways to fill the second cell with two balls, and after that we must place the remaining $n-2$ balls in $n-2$ slots, which can be done in $(n-2)!$ ways.

    Finally, we could have chosen any of the n slots to be empty, and any of the remaining $n-1$ slots to have 2 balls. So we end up with

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

19. a. A function of three variables can have up to

$$\binom{4+3-1}{4}$$

    fourth partial derivatives.

    b. Why? Well, to see how many rth order derivatives there are of a function with n variables, note that we are just putting r balls into n bins, or equivalently ordering r bars and $n-1$ stars. So we just need to choose r of the $r+n-1$ locations for bars, which can be done $\binom{r+n-1}{r}$ ways.

20. *12 phone calls come in at random on the 7 days of the week. What is the probability that no day goes by without a phone call?*

    First, we must dispel the idea that the answer can be computed in an unordered way. The ordering of the number of phone calls per day is important, since for example there are a lot more ways to get $(2, 2, 2, 2, 2, 1, 1)$ calls than there are to get $(12, 0, 0, 0, 0, 0, 0)$ calls – we would not want to undercount by only observing the number of calls per day. I arrived at a wrong answer using unordered counting, and thankfully there was an answer here to correct me.

In other words, we cannot place 12 things into 7 slots, since we will only count each arrangement instead of all of the ways it arises, decreasing its weight.

Rather, we choose 12 samples from a population of 7 with replacement and consider whether all 7 show up. That way, all of the different ways that calls can be assigned to days will be counted, rather than just counting the number of calls per day.

Indeed, 7^12 12 letter words can be made on 7 symbols. Some of those only include 6 or fewer symbols, so we need to subtract the count of words that include at most 6 of the 7 symbols. $\binom{7}{6}6^12$ will overcount those, since the each choice of 6 letters will double-count the words with 5 letters that can be made from those, so that all of the 5 letter words are counted twice, and so on for fewer letters, leading to the inclusion-exclusion

$$7^{12} - \binom{7}{6}6^{12} + \binom{7}{5}5^{12} - \binom{7}{4}4^{12} + \binom{7}{3}3^{12} - \binom{7}{2}2^{12} + \binom{7}{1}1^{12} = 3162075840.$$

In the end, we have $3162075840/7^12 \approx 0.22845$.

21. *A closet contains n pairs of shoes. If $2r < n$ shoes are chosen at random, what is the probability that there will be no matching pair in the sample?*

    First of all, there are $\binom{2n}{2r}$ total ways to choose $2r$ shoes. How many ways are there to choose without picking a pair? Well, we have to choose $2r$ of the n pairs, and then for each pair we pick one of the two shoes, for $\binom{n}{2r}2^{2r}$ ways.

24. b. *Two players A and B flip alternately and independently flip a biased coin, first player to obtain a heads wins, A flips first. What is the probability that A wins?

    A wins on the first turn w.p. p. A wins on the second turn with probability $p(1-p)^2$, third turn w.p. $p(1-p)^4$, etc.

    So, A wins with probability following the geometric series

$$\begin{aligned}\sum_{k=0}^{\infty} p(1-p)^{2k} &= p\sum_{k=0}^{\infty}((1-p)^2)^k \\ &= \frac{p}{1-(1-p)^2} \\ &= \frac{p}{1-(1-2p+p^2)} = \frac{1}{2-p}.\end{aligned}$$

    c. *Show that A wins with probability > 0.5 for all p.*

    Clearly the probability above is minimized at $p = 0$.

26. *A fair die is cast until a 6 appears. What is the probability that it must be cast more than five times?*

    This is $1 - \sum_{i=0}^{5} \frac{1}{6^i}$. I didn't have the formula for partial geometric series ready, but it turns out to be $\frac{1-1/6^n}{1-1/6}$, which makes sense.

27. *Verify the following identities for $n \geq 2$.*

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

       Indeed, for even n, this is clear since $\binom{n}{k} = \binom{n}{n-k}$.

       For odd n, let m such that $2m+1 = n$. Then the identity is

       $$\sum_{k=0}^{m}\binom{n}{2k} = \sum_{k=0}^{m}\binom{n}{2k+1}.$$

       But

       $$\binom{n}{2k} = \binom{n}{n-2k} = \binom{n}{2m+1-2k} = \binom{n}{2(m-k)+1},$$

       so that all the terms on the LHS appear on the RHS.

    b. Ibid.

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

       $$\begin{aligned}\sum_{k=1}^{n} k\binom{n}{k} &= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!} \\ &= \sum_{k=1}^{n} n\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} \\ &= n\sum_{k=0}^{n-1}\binom{n-1}{k-1},\end{aligned}$$

       and the result follows since

       $$\sum_{k=0}^{n-1}\binom{n-1}{k-1} = 2^{n-1},$$

       which is clear, since on the left we count the number of possible binary sequences of length $n-1$, which is also clearly counted on the right.

28. *Prove a weakening of Stirling's formula $n! \approx \sqrt{2\pi}n^{n+1/2}e^{-n}$, namely that*

    $$\lim_{n\to\infty}\frac{n!}{n^{n+1/2}e^{-n}}$$

    is some constant.

Indeed, following the hint, we note that since log is monotonically increasing, and since the average of a monotonic function on some interval has to be strictly greater than its right endpoint and less than its left endpoint, for all $k = 1, 2, \ldots$ we have

$$\int_{k-1}^{k} \log x \, dx < \log k < \int_{k}^{k+1} \log x \, dx.$$

Since

$$\log n! = \sum_{i=1}^{n} \log i,$$

it follows that

$$\begin{aligned} \int_0^n \log x \, dx &< \log n! < \int_1^{n+1} \log x \, dx \\ n \log n - n &< \log n! < (n+1)\log(n+1) - n \\ n \log n - n &< \log n! < (n+1)\log(n+1) - n \end{aligned}$$

The average of LHS and RHS is $\approx (n + 1/2)\log n - n$. At this point, we have good reason to believe that $\log n! - [(n+1/2)\log n - n]$ will converge to something, which is nice because we are interested in the limit of the exponential of this sequence. I'm going to leave this incomplete, though.

33. *5% of men and 25% of women are colorblind, and the world is half male and half female. What is $P(\text{male} \mid \text{colorblind})$?*

From this, we can see that $P(\text{colorblind}) = 0.15$. Then, by Bayes',

$$\begin{aligned} P(\text{male} \mid \text{colorblind}) &= \frac{P(\text{colorblind} \mid \text{male})P(\text{male})}{P(\text{colorblind})} \\ &= \frac{0.05 \cdot 0.5}{0.15} = \frac{1}{6}. \end{aligned}$$

In hindsight, since $P(\text{male}) = P(\text{female})$, we could have just looked at the ratios of the likelihoods.

34. (*Skipping the rodents.*) *We have two populations B, B, G and B, B, B, G, G. Let X be chosen by first picking a population at random and then picking an element at random from that population.*

a. $P(X = B) = \frac{1}{2}\frac{2}{3} + \frac{1}{2}\frac{3}{5} = \frac{19}{30}$.

b.
$$P(\text{pop. } 1 \mid X = B) = \frac{P(X = B \mid P1)P(P1)}{P(X = B)} = \frac{\frac{2}{3}\frac{1}{2}}{\frac{19}{30}} = \frac{10}{19}.$$

35. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $B \in \mathcal{F}$ a set with $P(B) > 0$. Then $P(\cdot \mid B)$ is a probability measure.*

First, let's show that $\mathcal{G} = \{A \cap B : A \in \mathcal{F}\}$ is a σ-algebra on the sample space B. Clearly $\varnothing \in \mathcal{G}$ since $\varnothing \in \mathcal{F}$. Further, if $A \in \mathcal{G}$, then $A \in \mathcal{F}$ and

$B \setminus A = B \cap A^C$ is clearly in $\mathcal{F}$ too, and since it is a subset of B it also lives in $\mathcal{G}$. Finally, since one cannot escape B with countable unions, $\mathcal{G}$ is also closed under countable union.

Clearly $P(B \mid B) = 1$ and nonnegativity is preserved. Finally, intersection with B preserves disjointness.

36. *Flip 10 coins with probability 1/5 of landing heads. What is the probability of at least 2 heads? What is the probability of at least two heads given at least one?*

    Well, the probability of at least one heads is

    $$1 - (\tfrac{4}{5})^1 0,$$

    two heads is

    $$1 - (\tfrac{4}{5})^1 0 - 10\frac{1}{5}(\tfrac{4}{5})^9,$$

    and the conditional probability is their quotient.

38. *If $P(B) = 1$, then $P(A \mid B) = P(A)$ for any A.*

    Indeed, write A as the disjoint union $A = (A \cap B) \cup (A \cap B^C)$, so that

    $$P(A) = P(A \cap B) + P(A \cap B^C).$$

    By definition,

    $$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = P(A \cap B).$$

    Now, assume for the sake of contradiction that $P(A \mid B) \equiv P(A \cap B) \neq P(A)$. By the first equation and positivity of measure, that would imply $P(A \cap B^C) > 0$. But again by disjoint additivity, that would imply $P(B^C) > 0$, which contradicts

    $$1 = P(\Omega) = P(B) + P(B^C).$$

    (b) and (c) are trivial. Instead of (d) as written, prove the chain rule for conditional probability: for any n sets $A_1, \ldots, A_n$ of positive measure,

    $$P(A_1 \cap \cdots \cap A_n) = P(A_n \mid A_{n-1} \cap \cdots \cap A_1) \ldots P(A_3 \mid A_2 \cap A_1) P(A_2 \mid A_1) P(A_1).$$

    The $n = 2$ case follows from the definition, and is the inductive step as well.

39. *Two events A and B can not be both mutually exclusive and independent.*

    (Note that if $P(A)$ or $P(B)$ is 0, sure, the statement can vacuously hold.)

    Let $P(A), P(B) > 0$. Then if $A \cap B = \varnothing$, clearly $P(A \cap B) \neq P(A)P(B)$. On the other hand, if $P(A \cap B) = P(A)P(B)$, then clearly the intersection cannot be empty.

41. *Consider a source that sends . and - in proportion 3:4, where errors can cause . to be received as - w.p. $\frac{1}{4}$ and - to be received as . w.p. $\frac{1}{3}$.*

a. *What is the probability that - was sent given that it was received?*

Well, we have $P(-\text{ sent}) = \frac{4}{7}$, $P(-\text{ recd} \mid -\text{ sent}) = \frac{2}{3}$, and

$$P(-\text{ recd}) = \frac{4}{7}\frac{2}{3} + \frac{3}{7}\frac{1}{4} = \frac{41}{84}.$$

Then,

$$\begin{aligned} P(-\text{ sent} \mid -\text{ recd}) &= \frac{P(-\text{ recd} \mid -\text{ sent})P(-\text{ sent})}{P(-\text{ recd})} \\ &= \frac{\frac{2}{3}\frac{4}{7}}{\frac{41}{84}} &= \frac{32}{41}. \end{aligned}$$

b. *Assuming independence and that .. was received, what is the probability of the four possible sent messages?*

$$P(..) = P(.\text{ sent} \mid .\text{ recd})^2,$$
$$P(-.) = P(.=) = P(.\text{ sent} \mid .\text{ recd})P(-\text{ sent} \mid .\text{ recd}),$$

and

$$P(--) = P(-\text{ sent} \mid .\text{ recd})^2.$$

Those can be calculated as in (a).

42. *Inclusion-exclusion.* Let $A_1, \ldots, A_n$ be events and consider $P(\cup_i = 1^n A_i)$. Let $A = \cup_{i=1}^n A_i$.

a. *Let E_k denote the set of sample points that are contained in exactly k of the events A_i. Then $P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(E_i)$.*

Indeed, by finite disjoint additivity of measure, it suffices to show that E_i form a disjoint partition of A.

First, clearly $A = \cup_{i=1}^n E_i$: every point in A is in some number k of events A_i, and thus is in E_k, and every point in some E_k must be in some events A_i.

Second, clearly the E_k are disjoint, since we used the word "exactly".

b. My text has a clear erratum here. And honestly it seems like (c) may as well, and I can't quite figure it out. Let's skip it.

46. *Seven balls are distributed randomly into seven cells. Let X_i be the number of cells containing i balls. Find the distribution of X_3.*

First of all, at most 2 cells can contain 3 balls. That will happen for permutations of the pattern 3310000, which can occur by: pick 2 locations for the 3s with $\binom{7}{2}$, pick three balls for the first three with $\binom{7}{3}$, pick three balls for the second three with $\binom{4}{3}$, and pick a location for the one with 5 for $\binom{7}{2}\binom{7}{3}\binom{4}{3}5 = 14700$.

To get $X_3 = 1$, what are the patterns, and how often do they occur?

Pattern	Count
3400000	$7\binom{7}{3}6 = 1470$
3220000	$7\binom{7}{3}\binom{6}{2}\binom{4}{2} = 22050$
3211000	$7\binom{7}{3}6\binom{4}{2}\binom{5}{2}2 = 176400$
3111100	$7\binom{7}{3}\binom{6}{4}4! = 88200$
Total	288120

So, since there are 7^7 possible arrangements, we have

$$P(X_3 = 2) = \frac{14700}{7^7} = \frac{300}{16807},$$
$$P(X_3 = 1) = \frac{288120}{7^7} = \frac{120}{343},$$
$$P(X_3 = 0) = 1 - \frac{300}{16807} - \frac{120}{343} = \frac{10627}{16807}.$$

47. *Prove that the following functions are cdfs.*

   a. $F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{atan} x$.

      Indeed, arctan is continuous with derivative $\frac{1}{1+x^2}$, so this function is increasing and continuous.

      What's more, $\operatorname{atan}(-\infty) = -\frac{\pi}{2}$ and $\operatorname{atan}(\infty) = \frac{\pi}{2}$, so we have correct limits of 0 and 1 for F.

   b. $G(x) = (1 + e^{-x})^{-1}$.

      Indeed, clearly this function is continuous, and it has derivative

      $$\frac{d}{dx}(1 + e^{-x})^{-1} = -(1 + e^{-x})^{-2}(-e^{-x}) = \frac{e^{-x}}{(1 + e^{-x})^2} > 0,$$

      so it's increasing. Finally, $e^{-\infty} = 0$ and $e^{\infty} \to \infty$, so that $G(-\infty) \to 0$ and $G(\infty) \to 1$.

48. *Let $F(x)$ be the cdf of a random variable X on $\mathbb{R}$. Then (a) $F(-\infty) = 0$ and $F(\infty) = 1$, (b) F is nondecreasing, and (c) F is right-continuous: i.e., for all x_0, $F(x+) = F(x_0)$.*

   Write $\mathbb{R} = \bigcup_{i \in \mathbb{Z}} [i, i+1)$. Then clearly

   $$F(-\infty) = \lim_{x \to -\infty} P(X \geq x) \leq \lim_{x \to 0} \sum_{i=-\infty}^{\lceil x+1 \rceil} P(X \in [i, i+1)).$$

   But if the RHS did not converge to 0, then $P(X \in \mathbb{R})$ could not be 1. Taking the limit in the other direction with the same argument leads to the sum $P(X \in \mathbb{R}) = 1$.

   F is nondecreasing since for all $x < y$, $P(X \leq x) \leq P(X \leq y)$.

F is right-continuous because we use the $\leq$ in the definition $F(x) = P(X \leq x)$. Indeed, take any x_0. Then for any $x > x_0$,

$$P(X \leq x) = P(X \leq x_0) + P(X \in (x_0, x]).$$

But $P(X \in (x_0, x]) \to 0$ as $x \to x_0$ by the axiom of continuity.

49. A cdf F_X is *stochastically greater than* a cdf F_Y if $F_X(t) \leq F_Y(t)$ for all t and there exists some t such that the inequality is strict. For such X and Y, show that (a) for all t,

$$P(X > t) \geq P(Y > t)$$

and (b) there is some t for which that inequality is strict.

a. Indeed, we have that for all t,

$$\begin{aligned} F_X(t) &\leq F_Y(t) \\ P(X \leq t) &\leq P(Y \leq t) \\ 1 - P(X > t) &\leq 1 - P(Y > t) \\ P(X > t) &\geq P(Y > t). \end{aligned}$$

b. Follows since we had some t where the first line in (a) was strict.

50. Verify the formula for partial sums of geometric series:

$$\sum_{k=0}^{n-1} r^k = \frac{1 - r^n}{1 - r}.$$

I don't like this formula, since I always forget, n or $n - 1$ or what?

First, consider $r \neq 1$, since there the sum is clearly just n.

We proceed by induction on n. For $n = 1$, the sum is just $r^0 = 1$, and the RHS is $\frac{1-r^1}{1-r} = 1$. To get a feel, for $n = 2$ the sum is $1 + r$, and the RHS is

$$\frac{1 - r^2}{1 - r} = \frac{(1 - r)(1 + r)}{1 - r} = (1 + r)\frac{1 - r}{1 - r} = 1 + r.$$

Now, assume that

$$\sum_{k=0}^{n-2} r^k = \frac{1 - r^{n-1}}{1 - r}.$$

Then it suffices to show that

$$r^{n-1} + \frac{1 - r^{n-1}}{1 - r} = \frac{1 - r^n}{1 - r}.$$

Indeed,

$$\begin{aligned} r^{n-1} + \frac{1-r^{n-1}}{1-r} &= \frac{(1-r)r^{n-1}}{1-r} + \frac{1-r^{n-1}}{1-r} \\ &= \frac{r^{n-1}-r^n}{1-r} + \frac{1-r^{n-1}}{1-r} \\ &= \frac{1-r^{n-1}+r^{n-1}-r^n}{1-r} \\ &= \frac{1-r^n}{1-r}, \end{aligned}$$

well ok then!

51. *An appliance store receives a shipment of 30 microwave ovens, 5 of which are defective. 4 are selected at random and tested. Let X be the number of defective ovens discovered. Calculate and plot its pmf and cdf.*

Let's do the calculation. Of all combinations of 4 ovens, there are $\binom{5}{4} = 5$ ways to choose four defective ones, $\binom{5}{3}\binom{25}{1} = 250$ ways to choose three, $\binom{5}{2}\binom{30}{2} = 3000$ for 2, you get the picture.

Here's a plot in Julia.

```
using Combinatorics, Plots
xs = [0, 1, 2, 3, 4]
ys = @. binomial(5, xs) * binomial(25, 4 - xs)
ys /= sum(ys)
plot(
    xs, cumsum(ys),
    linetype=:steppre, linecolor=:black, legend=false,
    size=(400, 400)
)
```

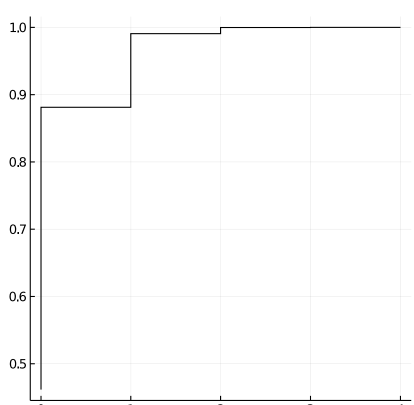

Figure 1.1: cdf of X.

52. *Let X be a continuous random variable with pdf f and cdf F. For some x_0 such*

that $F(x_0) < 1$, let

$$g(x) = \begin{cases} f(x)/[1 - F(x_0)] & x \geq x_0 \\ 0 & x < x_0. \end{cases}$$

Then $g(x)$ is a pdf.

Since g is clearly positive, it suffices to show that it integrates to 1, or in other words that $\int_{x_0}^{\infty} f = 1 - F(x_0)$. Indeed, by the definition of CDF and the FTC,

$$1 - F(x_0) = \lim_{x \to \infty} [F(x) - F(x_0)] = \int_{x_0}^{\infty} f(x)\, dx.$$

53. *A river floods every year. Suppose the low-water mark is at 1, and the high water mark Y has cdf*

$$F_Y(y) = 1 - \frac{1}{y^2}, \quad 1 \leq y < \infty.$$

   a. *Check that F_Y is a cdf.*

   This function is clearly continuous for $y \geq 1$. Further, differentiating, we get

$$f_Y(y) = \frac{2}{y^3},$$

   which is positive for positive y so that F_Y is increasing. Finally, $F_Y(1) = 0$ and $y^{-2} \to 0$ so that $F(\infty) = 1$.

   b. The pdf is $f_Y(y) = \frac{2}{y^3}$.

   c. *Let $Z = 10(Y - 1)$ and find F_Z.*

   Well,

$$\begin{aligned} F_Z(z) &= P(10(Y-1) \leq z) \\ &= P(10Y \leq z + 10) \\ &= P\Big(Y \leq \frac{z+10}{10}\Big) \\ &= F_Y\Big(\frac{z+10}{10}\Big) \\ &= 1 - \Big(\frac{z+10}{10}\Big)^{-2}, \end{aligned}$$

   on the domain $0 \leq z < \infty$.

54. *Find c such that the following $f(x)$ are pdfs.*

   a. $f(x) = c \sin x, 0 \leq x < \pi/2$.

   Pick

$$c = \int_0^{\frac{\pi}{2}} \sin x\, dx = [-\cos x]_0^{\frac{\pi}{2}} = -0 + 1,$$

   OK then, fine. Shoulda known.

b. $f(x) = ce^{-|x|}$, $x \in \mathbb{R}$.

By symmetry,

$$\frac{c}{2} = \int_0^\infty e^{-x}\,dx = [-e^{-x}]_0^\infty = -0 + \frac{1}{e},$$

so take $c = \frac{2}{e}$.

55. *An electronic device has lifetime T. The device has value $V = 5$ if it fails before time 3, otherwise it has value $V = 2T$. Find the cdf of V if T has pdf*

$$f_T(t) = \frac{1}{1.5}e^{-t/1.5}, \quad t > 0.$$

We can see that $V = 5$ with probability $P(T \leq 3)$, and otherwise V takes values starting at 6, such that

$$F_V(v) = \begin{cases} 0 & v < 5 \\ P(T \leq 3) & 5 \leq v < 6 \\ P(2T \leq v) & 6 \leq v. \end{cases}$$

These are more easily computed using the cdf of T, which is

$$F_T(t) = \int_0^t \tfrac{1}{1.5}e^{-s/1.5}\,ds = [-e^{-s/1.5}]_0^t = 1 - e^{-t/1.5}.$$

From here, we see that

$$P(T \leq 3) = F_T(3) = 1 - e^{-2},$$

and

$$P(2T \leq v) = F_T(\tfrac{v}{2}) = 1 - e^{-t/3}.$$

Chapter 2

Transformations and Expectations

1. *In each of the following find the pdf of Y and show that it integrates to 1.*

   a. $Y = X^3$ where $f_X(x) = 42x^5(1-x)$, $0 < x < 1$. We have

$$\begin{aligned}
f_Y(y) = \frac{d}{dy} F_Y(y) &= \frac{d}{dy} P(X \le \sqrt[3]{y}) \\
&= \frac{d}{dy} \int_0^{\sqrt[3]{y}} f_X(x)\, dx \\
&= f_X(\sqrt[3]{y}) \frac{1}{3} y^{-2/3} \\
&= \frac{42}{3} y^{5/3}(1 - y^{1/3}) y^{-2/3} \\
&= 14(y - y^{4/3})
\end{aligned}$$

   Integrating,

$$\int_0^1 y - y^{4/3}\, dy = \left[\frac{y^2}{2} - \frac{3}{7} y^{7/3} \right]_0^1 = \frac{1}{2} - \frac{3}{7} = \frac{1}{14}.$$

   b. $Y = 4X + 3$ with $f_X(x) = 7e^{-7x}$, $0 < x < \infty$.

   Here, $X = \frac{Y-3}{4}$. Above we did it by hand, but here we use the theorem

$$f_Y(x) = f_X(g^{-1}(y)) \left| (g^{-1})'(y) \right|.$$

   Note that here the domain changes, so $3 < y < \infty$ with pdf

$$f_Y(y) = 7e^{-7(y-3)/4} = \frac{7}{4} e^{21/4} e^{-7y/4}.$$

Integrating,

$$\int_3^\infty e^{-7y/4}\,dx = \left[-\frac{4}{7}e^{-7y/4}\right]_3^\infty = \frac{4}{7}e^{-21/4}.$$

c. $Y = X^2$ and $f_X(x) = \frac{15}{2}x^2(1-x^2), 0 < x < 1$.

(There is a problem in the question as stated, f_X does not integrate to 1. So it's modified here.)

Now, $X = \sqrt{Y}$ and the domain does not change. Applying the theorem,

$$f_Y(y) = \frac{15}{2}(x - x^2)\,\frac{1}{2\sqrt{x}} = \frac{15}{4}(\sqrt{x} - x^{3/2}).$$

Integrating,

$$\int_0^1 \sqrt{x} - x^{3/2}\,dx = \left[\frac{2}{3}x^{3/2} - \frac{2}{5}x^{5/2}\right]_0^1 = \frac{2}{3} - \frac{2}{5} = \frac{4}{15}.$$

2. *Find the pdf of Y.*

a. $Y = X^2, f_X(x) = 1, 0 < x < 1$.

The domain does not change, and

$$f_Y(y) = \frac{1}{2\sqrt{y}}.$$

b. $Y = -\log X$, where X has pdf

$$f_X(x) = \frac{(n+m+1)!}{n!m!}x^n(1-x^m)$$

on $0 < x < 1$.

$-\log$ is decreasing, so the domain becomes $0 < y < \infty$ since $X = e^{-Y}$. We have

$$\left|\frac{d}{dy}g^{-1}(y)\right| = e^{-y},$$

so that

$$f_Y(y) = \frac{(n+m+1)!}{n!m!}e^{-(n+1)y}(1-e^{-my}).$$

Looks familiar...

c. $Y = e^X$ and X has pdf

$$f_X(x) = \frac{1}{\sigma^2}xe^{-(x/\sigma)^2/2}.$$

Then $X = \log Y$ and the domain does not change. The pdf becomes

$$f_Y(y) = \frac{1}{\sigma^2}e^{-\log(x)^2/2\sigma^2}.$$

3. Let $X \sim \text{Geom}(\frac{1}{3})$, so that it has pmf

$$\frac{1}{3}\left(\frac{2}{3}\right)^x, \quad \text{x=0,1,2,...}$$

Let $Y = \frac{X}{X+1}$ and determine its distribution.

The function
$$g(x) = \frac{x}{x+1}$$
maps $0, 1, \ldots$ onto $0, \frac{1}{2}, \frac{2}{3}, \ldots$ with the same probability as x. If

$$\begin{aligned} y &= \frac{x}{x+1} \\ (x+1)y &= x \\ xy + y &= x \\ y &= x - xy \\ y &= x(1-y) \\ x &= \frac{y}{1-y}, \end{aligned}$$

then

$$f_Y(y) = P(\frac{X}{X+1} = y) = P(X = \frac{y}{1-y}) = \frac{1}{3}\left(\frac{2}{3}\right)^{y/(1-y)}.$$

4. Let $f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}$ for $\lambda > 0, x \in \mathbb{R}$.

   a. *Verify that $f(x)$ is a pdf.*

   First, $f > 0$ for all x and is clearly continuous. Second,

$$\begin{aligned} \int_{\mathbb{R}} f(x)\, dx &= \lambda \int_0^\infty e^{-\lambda x}\, dx \\ &= \lambda\left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^\infty \quad = \lambda(\frac{1}{\lambda}) = 1. \end{aligned}$$

   b. We have

$$P(X < t) = \begin{cases} \frac{\lambda}{2}\int_{-\infty}^t e^{\lambda x}\, dx & \text{if} \quad t < 0 \\ \frac{\lambda}{2}\left(\frac{1}{2} + \int_0^t e^{-\lambda x}\, dx\right) & \text{if} \quad t \geq 0. \end{cases}$$

Evaluating these integrals, we get

$$\begin{aligned}\int_{-\infty}^{t} e^{\lambda x}\,dx &= \left[\frac{1}{\lambda}e^{\lambda x}\right]_{-\infty}^{t} \\ &= \frac{1}{\lambda}e^{\lambda t}, \quad \text{and} \\ \int_{0}^{t} e^{-\lambda x}\,dx &= \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_{0}^{t} \\ &= -\frac{1}{\lambda}e^{-\lambda t} + \frac{1}{\lambda}.\end{aligned}$$

Plugging back in,

$$P(X < t) = \begin{cases} \frac{1}{2}e^{\lambda t} & \text{if} \quad t < 0 \\ \frac{1}{2}\frac{\lambda}{2}\left(\frac{1}{2} + \int_0^t e^{-\lambda x}\,dx\right) & \text{if} \quad t \geq 0. \end{cases}$$

c. By symmetry,

$$\begin{aligned}P(|X| < t) &= 2\int_0^t \frac{\lambda}{2}e^{-\lambda x}\,dx \\ &= \lambda\left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^t \\ &= 1 - e^{-\lambda t}.\end{aligned}$$

5. *Find the pdf of* $Y = \sin^2(x)$ *where* X *is uniformly distributed on* $(0, 2\pi)$.

   First find it by differentiating the cdf given in $(2.1.6)$. To do so, we need to find two solutions to $\sin^2(x) = y$ in $(0, \pi)$. We can take

$$\begin{aligned}x_1 &= \operatorname{asin}\sqrt{y} \\ x_1 &= \pi + \operatorname{asin}\left(-\sqrt{y}\right).\end{aligned}$$

   Thus we need to differentiate

$$\begin{aligned}P(Y \leq y) &= 2P(X \leq \operatorname{asin}\sqrt{y}) + 2P(\pi + \operatorname{asin}\left(-\sqrt{y}\right) \leq X \leq \pi) \\ &= \frac{1}{\pi}\operatorname{asin}\sqrt{y} + \frac{1}{\pi}\left[-\operatorname{asin}\left(-\sqrt{y}\right)\right] \\ &= \frac{2}{\pi}\operatorname{asin}\sqrt{y}.\end{aligned}$$

   So, what's obvious now is that there's another symmetry, so that the probability is four times the probability in the first region. Also, this is clearly a cdf, since for $0 \leq y \leq 1$ this function's range is $[0, 1]$. Differentiating,

$$\begin{aligned} f_Y(y) &= \frac{2}{\pi}\frac{d}{dy}\operatorname{asin}\sqrt{y} \\ &= \frac{2}{\pi}\frac{1}{\sqrt{1-y}}\frac{1}{2\sqrt{y}} \\ &= \frac{1}{\pi}\frac{1}{\sqrt{y-y^2}}. \end{aligned}$$

I don't want to find this using Theorem 2.1.8 since Theorem 2.1.8 would lose all of the symmetries.

6. *Find the pdf of Y and show that it integrates to 1.*

a. $f_X(x) = \frac{1}{2}e^{-|x|}$, $x \in \mathbb{R}$, $Y = |X|^3$.

Here,

$$\begin{aligned} P(Y \le y) &= P(-\sqrt[3]{y} \le X \le \sqrt[3]{y}) \\ &= \frac{1}{2}\int_{-\sqrt[3]{y}}^{\sqrt[3]{y}} e^{-|x|}\,dx \\ &= \int_0^{\sqrt[3]{y}} e^{-x}\,dx. \end{aligned}$$

Differentiating,

$$f_Y(y) = \frac{1}{3}y^{-2/3}e^{-\sqrt[3]{y}}.$$

To show that this integrates to 1, make the change of variable $s = \sqrt[3]{y}$, $ds = \frac{1}{3}y^{-2/3}\,dy$. Note this does not change the domain.

$$\int_0^\infty f_Y(y)\,dy = \int_0^\infty e^{-s}\,ds = [-e^{-s}]_0^\infty = 0 + 1.$$

b. $f_X(x) = \frac{3}{8}(x+1)^2$, $-1 < x < 1$, $Y = 1 - X^2$.

Here, $0 < 1 - x^2 < 1$ is the range of Y. Picturing $g(x) = 1 - x^2$ as an inverted frown, since it has inverses $g^{-1}(y) = \pm\sqrt{1-y}$, we can see that

$$\begin{aligned} P(Y \le y) &= P(-1 \le X \le -\sqrt{1-y}) + P(\sqrt{1-y} \le X \le 1) \\ &= \int_{-1}^{-\sqrt{1-y}} \frac{3}{8}(x+1)^2\,dx + \int_{\sqrt{1-y}}^{1} \frac{3}{8}(x+1)^2\,dx. \end{aligned}$$

Differentiating,

$$\begin{aligned} f_Y(y) &= \frac{3}{8}(1-\sqrt{1-y})^2\frac{1}{2\sqrt{1-y}} + \frac{3}{8}(\sqrt{1-y}+1)^2\frac{1}{2\sqrt{1-y}} \\ &= \frac{3}{16}\frac{(1-\sqrt{1-y})^2 + (1+\sqrt{1-y})^2}{\sqrt{1-y}} \\ &= \frac{3}{8}(\sqrt{1-y}^{-1} + \sqrt{1-y}). \end{aligned}$$

Integrating again,

$$\frac{3}{8}\int_0^1 \sqrt{1-y}^{-1} + \sqrt{1-y}\,dy = \frac{3}{8}\left[-2\sqrt{1-y} - \frac{2}{3}(1-y)^{3/2}\right]_0^1$$
$$= \frac{3}{8}\left[2 + \frac{2}{3}\right] = 1.$$

c. $f_X(x) = \frac{3}{8}(x+1)^2, -1 < x < 1,$

$$Y = \begin{cases} 1 - X^2 & X \leq 0 \\ 1 - X & X > 0. \end{cases}$$

Again, the range is the same, but this one is just more complicated. Now, we have

$$\begin{aligned} P(Y \leq y) &= P(-1 \leq X \leq -\sqrt{1-y}) + P(1-y \leq X \leq 1) \\ &= \int_{-1}^{-\sqrt{1-y}} \frac{3}{8}(x+1)^2\,dx + \int_{1-y}^{1} \frac{3}{8}(x+1)^2\,dx. \end{aligned}$$

Then

$$\begin{aligned} f_Y(y) &= \frac{3}{16}\frac{(1-\sqrt{1-y})^2}{\sqrt{1-y}} + \frac{d}{dy}\int_{1-y}^{1} \frac{3}{8}(x+1)^2\,dx \\ &= \frac{3}{16}\left(\sqrt{1-y}^{-1} - 2 + \sqrt{1-y}\right) + \frac{3}{8}(2-y)^2. \end{aligned}$$

Integrating (this will be fun),

$$\begin{aligned} \int_0^1 f_Y(y)\,dy &= \frac{3}{16}\int_0^1 \sqrt{1-y}^{-1} - 2 + \sqrt{1-y}\,dy + \frac{3}{8}\int_0^1 (2-y)^2 \\ &= 1. \end{aligned}$$

7. *Let X have pdf $f_X(x) = \frac{2}{9}(x+1)$ on $-1 \leq x \leq 2$.*

a. *Find the pdf of $Y = X^2$.*

Well,

$$\begin{aligned} P(Y \leq y) &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} \mathbb{1}_{-1 \leq x \leq 2} f_X(x)\,dx. \end{aligned}$$

Differentiating,

$$\begin{aligned} f_Y(y) &= \frac{2}{9}\frac{1}{2y^{1/2}}\left[\mathbb{1}_{-1 \leq \sqrt{y} \leq 2}(\sqrt{y}+1) - \mathbb{1}_{-1 \leq -\sqrt{y} \leq 2}(1-\sqrt{y})\right] \\ &= \frac{1}{9y^{1/2}}\left[\mathbb{1}_{0 \leq y \leq 1}(\sqrt{y} + 1 - 1 + \sqrt{y}) + \mathbb{1}_{1 \leq y \leq 4}(\sqrt{y}+1)\right] \\ &= \frac{1}{9}\left[2\mathbb{1}_{0 \leq y \leq 1} + \mathbb{1}_{1 \leq y \leq 4}(1 + y^{-1/2}).\right] \end{aligned}$$

Clearly this is positive, but check that it integrates to 1:

$$\begin{aligned}\frac{9}{1}\int_{-1}^{4} f_Y(y)\,dy &= 2\int_0^1 2\,dy + \int_1^4 1 + y^{-1/2}\,dy \\ &= 2[2y]_0^1 + \left[y + 2\sqrt{y}\right]_1^4 \\ &= 4 + [(4+4) - (1+2)] = 9.\end{aligned}$$

b. Theorem 2.1.8 can be extended by just assuming that $f_X = 0$ away from its support. Then one can apply it to this example just fine. I don't like calculating with that cumbersome theorem though, the math that results is identical to the above but with more hoops to jump through.

8. *In each of the following, show that F_X is a cdf and find the quantile function F_X^{-1}.*

a.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x} & x \geq 0. \end{cases}$$

Clearly $F_X(-\infty) = 0$ and $F_X(\infty) = 1 - e^{-\infty} = 1$. Also, this function is continuous, since $1 - e^{-0} = 0$, and clearly it's increasing.

Clearly $F_X^{-1}(0) = \infty$, but for $x > 0$, F_X is invertible. Since

$$y = 1 - e^{-x} \implies 1 - y = e^{-x} \implies -x = \log(1-y),$$

we have

$$F_X^{-1}(y) = -\log(1-y).$$

b.

$$F_X(x) = \begin{cases} e^x/2 & x < 0 \\ 1/2 & 0 \leq x < 1 \\ 1 - e^{1-x}/2 & 1 \leq x. \end{cases}$$

Notice that $e^{-\infty}/2 = 0 = e^{1-\infty}/2$, so that $F_X(-\infty) = 0$ and $F_X(\infty) = 1$. Also, $e^0/2 = 1/2 = 1 - e^{1-1}/2$, so that F_X is continuous. Finally, the function is increasing.

For the quantile function, we can invert the cdf for $y < \frac{1}{2}$ and $y > \frac{1}{2}$, and use the infimum for $y = \frac{1}{2}$, so that

$$F_X^{-1}(y) = \begin{cases} \log 2y & y < \frac{1}{2} \\ 0 & y = \frac{1}{2} \\ 1 - \log(2 - 2y) & y > \frac{1}{2}. \end{cases}$$

c.

$$F_X(x) = \begin{cases} e^x/4 & x < 0 \\ 1 - e^{-x}/4 & x \geq 0. \end{cases}$$

F_X has a jump discontinuity at 0, since $e^0/4 = \frac{1}{4}$ and $1 - e^0/4 = \frac{3}{4}$. However, it is continuous from the right with limits from the left since we choose the branch to the right at 0. Still, it's increasing and has the right limits since $e^{-\infty} = 0$.

For the quantile function, F_X is injective so it can be inverted without needing to take infima. The only difficulty is to recognize that for $\frac{1}{4} \leq y < \frac{3}{4}$, $F_X(x) \geq y$ implies that we use the branch for $x \geq 0$, so that

$$F_X^{-1}(y) = \begin{cases} \log 4y & y < \frac{1}{4} \\ -\log(4 - 4y) & y \geq \frac{1}{4}. \end{cases}$$

9. *If the random variable X has pdf $f(x) = \mathbb{1}_{1<x<3} \frac{x-1}{2}$, find a monotone function $u(x)$ such that $Y = u(X)$ has the uniform(0,1)\$ distribution.*

   By the probability integral transformation, we can take $u = F_X$. Indeed,

$$F_X(x) = \int_{-\infty}^{x} \frac{s-1}{2}\, ds = \begin{cases} 0 & x \leq 1 \\ \frac{x^2}{4} - \frac{x}{2} - \frac{1}{4} & 1 < x < 3 \\ 1 & x \geq 3. \end{cases}$$

   Since $x \leq 1$ or $x \geq 3$ w.p. 0, we can take $u(x) = \frac{x^2}{4} - \frac{x}{2} - \frac{1}{4}$.

10. *Let X be discrete with cdf F_X and let $Y = F_X(X)$.*

    a. *Prove that Y is stochastically greater than $U \sim$ uniform$(0, 1)$.*

    It suffices to show that for all y,

$$P(Y \leq y) \leq P(U \leq y) = y,$$

    and that there exists some y such that the inequality is strict.

    Indeed, let $\mathcal{X} = \{x : P(X = x) > 0\}$. Then for each y,

$$P(Y \leq y) = \sup_{x \in X : F_X(x) \leq y} F_X(x) \leq y.$$

    Thus we have the non-strict inequality, and in fact if there exists $x \in \mathcal{X}$ such that $F_X(x) = y$, then the supremum is achieved and $P(Y \leq y) = y$. Now, assume for the sake of contradiction that the supremum is achieved anywhere. Then for all y, $P(Y \leq y) = y$. But then Y is a continuous random variable, and $P(Y = y) = 0$ for all y. But that contradicts the assumption that X was discrete, since if X is discrete, there exists some $x \in \mathcal{X}$ such that $P(Y = F(X_i)) = P(X = x_i) > 0$.

    b. *This asks for another proof of the same claim, but to provide the y such that the inequality is strict constructively. I think there's a typo, I can't see how to do it by adding ϵ, only by subtracting ϵ (going to the left as below).*

To find a y such that the inequality is strict, take arbitrary $x \in \mathcal{X}$, and let $p = P(X = x)$. We claim that there is no x' such that $F_X(x) - p < F(x') < F_X(x)$. Indeed, if there were such x', then

$$P(X = x) \leq P(x' < X \leq x) = F_X(x) - F(x') < F_X(x) - F_X(x) - p = p,$$

implying that $P(X = x) = p < p$.

Now, let y such that $F_X(x) - p < y < F_X(x)$. Then

$$P(Y \leq y) = \sup_{x \in X : F_X(x) \leq y} F_X(x) \leq F(x) - p < y.$$

11. *Let $X \sim N(0, 1)$.*

a. *Find EX^2 directly, and then by using the χ_1^2 pdf.*

After going down a lot of wrong roads, we can integrate by parts using $u = x$ and $dv = xe^{-x^2/2}\,dx$, so that

$$v = \int xe^{-x^2/2}\,dx = \int -e^{-s}\,ds = -e^{-x^2/2}.$$

Then

$$\int_{-\infty}^{\infty} x\,xe^{-x^2/2}\,dx = \left[-xe^{-x^2/2}\right]_{-\infty}^{\infty} + \int -e^{-x^2/2}\,dx \quad = 0 + \sqrt{2\pi},$$

which we know by recognizing the kernel of the Gaussian pdf, so that the result is 1 as it should be.

Now, let $Y = X^2$. Then

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sqrt{y}}(f_X(-\sqrt{(}y)) + f_X(\sqrt{y})) \\ &= \frac{1}{2\sqrt{y}}\frac{1}{\sqrt{2\pi}}2e^{-y/2} \\ &= \frac{1}{\sqrt{2\pi y}}e^{-y/2}. \end{aligned}$$

Then we just want the expectation

$$\begin{aligned} EY &= \frac{1}{\sqrt{2\pi}}\int_0^{\infty}\sqrt{y}e^{-y/2}\,dy \\ &= \frac{1}{\sqrt{2\pi}}\left(\left[-2\sqrt{y}e^{-y/2}\right]_0^{\infty} + 2\int_0^{\infty}\frac{1}{2\sqrt{y}}e^{-y/2}\,dy\right) \\ &= \frac{1}{\sqrt{2\pi}}\int_0^{\infty}\frac{1}{\sqrt{y}}e^{-y/2}\,dy \\ &= \frac{\sqrt{2\pi}}{\sqrt{2\pi}} = 1. \end{aligned}$$

b. *Find the pdf of $Y = |X|$ and compute the mean and variance.*

To compute the pdf, we have

$$f_Y(y) = \frac{d}{dy} 2F_X(y) = \sqrt{\tfrac{2}{\pi}} e^{-y^2/2}.$$

Then the expectation is

$$EY = \sqrt{\tfrac{2}{\pi}} \int_0^\infty y e^{-y^2/2}\, dy = \sqrt{\tfrac{2}{\pi}} [-e^{-y^2/2}]_0^\infty = \frac{1}{\sqrt{2\pi}}.$$

And the variance can be computed with $\operatorname{var} Y = E[Y^2] - E[Y]^2$. Luckily, $Y^2 = X^2$, so we know that

$$\operatorname{var} Y = 1 - \frac{1}{2\pi}.$$

13. *Let X be the length of the run started by the first of a sequence Y_i of independent coin flips. Find the distribution of X and EX.*

Conditioned on the value of the first flip, X is geometrically distributed:

$$P(X = k \mid Y_1 = y) = 0.5^{k-1} 0.5.$$

Thus

$$P(X = k) = P(X = k \mid Y_1 = y) P(Y_1 = y) = 0.5^k 0.5$$

is also geometrically distributed (k failures parametrization), and has expectation $EX = \frac{1-p}{p} = 1$.

14. *Let X have pdf f such that $f(x) = 0$ for $x < 0$. Show the "bathtub formula"*

$$EX = \int_0^\infty 1 - F(x)\, dx.$$

There is an easy visual proof: imagine filling the area between $F(x)$ and 1 by a Riemann-style sum, only with the rectangles based on the y-axis. Then they have height $F(dx) = dP(x) = f(x)\, dx$ and width x, so we are summing $x\, dP = x f(x)\, dx$.

We can translate this into words using Fubini's theorem to interchange integration from x to y:

$$\begin{aligned}
\int_0^\infty 1 - F(x)\, dx &= \int_0^\infty P(X > x)\, dx \\
&= \int_0^\infty E[\mathbb{1}_{X>x}]\, dx \\
&= \int_0^\infty \int_0^\infty \mathbb{1}_{y>x} f(y) dy\, dx \\
&= \int_0^\infty \left[\int_0^\infty \mathbb{1}_{y>x}\, dx \right] f(y)\, dy \\
&= \int_0^\infty y f(y)\, dy.
\end{aligned}$$

15. *Let X and Y be random variables and* $X \wedge Y = \min(X, Y)$, $X \vee Y = \max(X, Y)$. *Then*

$$E[X \vee Y] = EX + EY - E[X \wedge Y].$$

Indeed, notice that for all ω,

$$(X \wedge Y + X \vee Y)(\omega) = \begin{cases} Y(\omega) + X(\omega) & \text{if} \quad X(\omega) > Y(\omega) \\ X(\omega) + Y(\omega) & \text{if} \quad X(\omega) \le Y(\omega) \end{cases} = (X + Y)(\omega).$$

Thus, taking expectations of both sides,

$$E[X \wedge Y + X \vee Y] = E[X + Y]$$

and, by linearity,

$$E[X \wedge Y] + E[X \vee Y] = EX + EY.$$

Subtract $E[X \wedge Y]$ to finish.

This requires that these be defined on the same probability space. But actually, I think that's sort of an assumption of the construction of taking min and max, since events like $\{X > Y\}$ need to be measurable.

16. *Let T have survival function* $P(T > t) = ae^{-\lambda t} + (1 - a)e^{-\mu t}$ *for* $\lambda, \mu > 0$, $0 < a < 1$, *and* $t > 0$. *Compute ET.*

By a previous exercise, since T is positive,

$$\begin{aligned} ET &= \int_0^\infty ae^{-\lambda t} + (1 - a)e^{-\mu t}\, dt \\ &= \left[-\frac{a}{\lambda} e^{-\lambda t} - \frac{1 - a}{\mu} e^{-\mu t} \right]_0^\infty \\ &= \frac{a}{\lambda} + \frac{1 - a}{\mu}. \end{aligned}$$

17. *Find the median of the following distributions.*

    a. $f(x) = 3x^2$, $0 < x < 1$.

    Since f is strictly positive, the median is the unique m such that

    $$\int_0^m 3x^2\, dx = \frac{1}{2}.$$

    Indeed,

    $$\int_0^m 3x^2\, dx = \left[x^3\right]_0^m = m^3,$$

    so take $m = \sqrt[3]{\frac{1}{2}}$.

    b. $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, $-\infty < x < \infty$.

    The Cauchy is symmetric, so clearly $m = 0$.

18. *Show that any median m of a continuous random variable minimizes the mean absolute deviation, i.e. if m is a median of X with pdf f, then*

$$\min_a E|X-a| = E|X-m|.$$

For any a, since $X = a$ with probability 0, we can break up the following integral into two integrals of differentiable functions to allow for the use of Leibniz's rule (although I think it should apply up to a set of measure 0 anyway). Indeed,

$$\begin{aligned}
\frac{d}{da}E|X-a| &= \frac{d}{da}\int_{-\infty}^{\infty}|x-a|f(x)\,dx \\
&= \frac{d}{da}\left[\int_{-\infty}^{a}(a-x)f(x)\,dx + \int_{a}^{\infty}(x-a)f(x)\,dx\right] \\
&= \left[(a-a)f(a) + \int_{-\infty}^{a}f(x)\,dx\right] + \left[-(a-a)f(a) - \int_{a}^{\infty}f(x)\,dx\right] \\
&= \int_{-\infty}^{a}f(x)\,dx - \int_{a}^{\infty}f(x)\,dx.
\end{aligned}$$

Thust to achieve a minimum, we need

$$\int_{-\infty}^{a}f(x)\,dx = \int_{a}^{\infty}f(x)\,dx = \frac{1}{2},$$

so that a is a median.

19. *Show that the minimizer a of $E(X-a)^2$ is the mean EX, and list the necessary assumptions on F and f.*

Indeed, by Leibniz's rule,

$$\begin{aligned}
\frac{d}{da}E(X-a)^2 &= \frac{d}{da}\int_{-\infty}^{\infty}(x-a)^2 f(x)\,dx \\
&= -2\int_{-\infty}^{\infty}(x-a)f(x)\,dx. \quad = -2\left[\int_{-\infty}^{\infty}xf(x)\,dx - a\int_{-\infty}f(x)\,dx\right] = -2[EX -
\end{aligned}$$

Setting this equal to 0, we see that a is a critical point iff it is the mean.

To see that the critical point is a minimum, differentiate again to obtain

$$\frac{d}{da}(-2[EX-a]) = 2,$$

so that the function is convex.

The only assumption was the existence of absolute continuity (the same proof would work with a pmf), but I'm sure that can't be necessary...

20. *A couple decides to flip a coin until it lands heads. What is the expected total number of coin tosses?*

Clearly the pmf is $P(X=k) = 0.5^{k-1}0.5$, the $k-1$ failures parametrization of the geometric distribution, which has expectation $\frac{1}{p}$.

21. *Prove the "two-way rule" for expectations, that $EY = Eg(X)$ if $Y = g(X)$.*

I am a little confused why they would ask this, $Y = g(X)$ with probability 1 so of course they have the same expectation. They allow us to assume that g is monotonic, so my guess is they want us to show that

$$\int_{\mathcal{X}} x f_X(x)\,dx = \int_{g(\mathcal{X})} g(x) f_X(g^{-1}(x)) \frac{1}{|g'(x)|}\,dx = \int_{\mathcal{Y}} y f_Y(y)\,dy,$$

but here the first equality is just a change of variables and the second is the expectation of Y. It remains to observe that the density of y is as given in the middle, but this has been shown already...

22. *Let X have pdf*

$$f(x) = \frac{4}{\beta^3\sqrt{\pi}} x^2 e^{-x^2/\beta^2}$$

for $0 < x < \infty$ and $\beta > 0$.

a. *Verify that f is a pdf.*

Indeed, f is positive, so it remains to integrate it. But

$$\begin{aligned}
\int_0^\infty x^2 e^{-x^2/\beta^2} &= \frac{1}{2}\int_{-\infty}^{\infty} x^2 \exp\left(-\frac{x^2}{2(\beta/\sqrt{2})^2}\right) dx \\
&= \frac{\sqrt{\pi}\beta}{2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}(\beta/\sqrt{2})} x^2 \exp\left(-\frac{x^2}{2(\beta/\sqrt{2})^2}\right) dx \\
&= \frac{\sqrt{\pi}\beta}{2} E[N(0,\beta^2/2)^2] \\
&= \frac{\sqrt{\pi}\beta}{2}\frac{\beta^2}{2} \\
&= \frac{\sqrt{\pi}\beta^3}{4},
\end{aligned}$$

where we have recognized the variance of a Gaussian distribution to make the computation easier.

b. This is a lot of integration by parts...

23. Let X have pdf $f(x) = (1+x)/2$, $-1 < x < 1$.

a. *Find the pdf of $Y = X^2$.*

For $0 \leq y < 1$,

$$\begin{aligned}
F_Y(y) &= \frac{d}{dy}\int_{-\sqrt{y}}^{\sqrt{y}} f(x)\,dx \\
&= (f(\sqrt{y})\frac{1}{2\sqrt{y}} - f(-\sqrt{y})\frac{-1}{2\sqrt{y}}) \\
&= \frac{1+\sqrt{y}+1-\sqrt{y}}{4\sqrt{y}} \\
&= \frac{1}{2\sqrt{y}}.
\end{aligned}$$

Check that it's a pdf:

$$\int_0^1 \frac{1}{2\sqrt{y}}\,dy = \left[\sqrt{y}\right]_0^1 = 1.$$

b. *Find EY and* var Y.

$$EY = \int_0^1 \frac{\sqrt{y}}{2}\,dy = \left[\frac{1y^{3/2}}{3}\right]_0^1 = \frac{1}{3}.$$

And, since

$$EY^2 = \int_0^1 \frac{y^{3/2}}{2}\,dy = \left[\frac{1y^{5/2}}{5}\right]_0^1 = \frac{1}{5},$$

we have $\operatorname{var} Y = E[Y^2] - E[Y]^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$.

24. *Compute EX and* var X *for each of the following distributions.*

a. $f(x) = ax^{a-1}$, $0 < x < 1$, $a > 0$.

$$\begin{aligned}
EX &= \int_0^1 ax^a\,dx = \left[\frac{a}{a+1}x^{a+1}\right]_0^1 = \frac{a}{a+1}, \\
EX^2 &= \int_0^1 ax^{a+1}\,dx = \left[\frac{a}{a+2}x^{a+2}\right]_0^1 = \frac{a}{a+2}, \\
\operatorname{var} X &= E[X^2] - E[X]^2 = \frac{a}{a+2} - \frac{a^2}{(a+1)^2}.
\end{aligned}$$

b. $f(x) = \frac{1}{n}$, $x = 1, 2, \ldots, n$.

Then

$$
\begin{aligned}
EX &= \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}, \\
EX^2 &= \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}, \\
\operatorname{var} X &= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} \\
&= \frac{(8n^2+12n+4)-(6n^2+12n+6)}{24} = \frac{n^2-1}{12}.
\end{aligned}
$$

c. $f(x) = \frac{3}{2}(x-1)^2, 0 < x < 2.$

To simplify, let $Y = X - 1$ with pdf $g(y)$, so that

$$g(y) = \frac{3}{2}y^2$$

on $-1 < y < 1$. Then by symmetry, $EY = 0$, so that $EX = 1$.

Now, notice that $Y = X - EX$, so that $EY^2 = E[(X-EX)^2] = \operatorname{var} X$. So,

$$\operatorname{var} X = \int_{-1}^{1} \frac{3}{2}y^4\,dy = \left[\frac{3}{10}y^5\right]_{-1}^{1} = \frac{3}{5}.$$

25. *Let X have pdf f, where f is even. Then, ...*

a. *X and $-X$ are identically distributed.*

Indeed, since $\left|\frac{d}{dx} - x\right| = 1, f_{-X}(x) = f(-x)$.

b. *$M_X(t)$ is even.*

Notice that making the change of variables $u = -x$, $du = -dx$,

$$
\begin{aligned}
M_X(-t) &= Ee^{-tX} \\
&= \int_{-\infty}^{\infty} e^{-tx} f(x)\,dx \\
&= -\int_{\infty}^{-\infty} e^{tx} f(-x)\,dx \\
&= \int_{-\infty}^{\infty} e^{tx} f(-x)\,dx \\
&= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx \\
&= M_X(t).
\end{aligned}
$$

26. *A pdf $f(x)$ is said to be symmetric about a if $f(a+x) = f(a-x)$ for all x.*

a. *Give three examples of symmetric pdfs.*

$N(a,1)$: $f(x) = \frac{1}{\sqrt{2\pi}}e^{-(x-a)^2/2}$.

Cauchy centered at a: $f(x) = \frac{1}{\pi}\frac{1}{1+(x-a)^2}$.

Double exponential centered at a: $f(x) = \frac{1}{2}e^{-|x-a|}$.

b. *Let $X \sim f(x)$ symmetric about a. Then the median of X is a.*

Indeed, since f is symmetric,

$$\int_{-\infty}^{a} f(x)\,dx = \int_{a}^{\infty} f(x)\,dx.$$

c. *If f is symmetric and EX exists, then $EX = a$.*

Indeed, let $Y = X - a$ so that $EY = EX - a$. Then

$$P(Y \le y) = P(X - a \le y) = P(X \le y + a),$$

so that $f_Y(y) = f(y+a)$. Then

$$\int_{-\infty}^{\infty} y f_Y(y)\,dy = \int_{-\infty}^{\infty} y f(y+a)\,dy \quad = \int_{0}^{\infty} (-yf(-y+a) + yf(y+a))\,dy = 0.$$

d. *Show that $f(x) = \mathbb{1}_{x\ge 0}\, e^{-x}$ is not symmetric.*

Indeed, assume that f were symmetric about a. But then $f(a+2|a|) > 0$ while $f(a-2|a|) = 0$.

e. *Show that the median is less than the mean for that f.*

Indeed, the median satisfies

$$\begin{aligned}\int_0^a e^{-x}\,dx &= \int_a^\infty e^{-x}\\ [-e^{-x}]_0^a &= [-e^{-x}]_a^\infty\\ 1 - e^{-a} &= e^{-a},\end{aligned}$$

so that $2e^{-a} = 1$, or $a = \log 2$.

But the mean is

$$\begin{aligned}\int_0^\infty xe^{-x}\,dx &= [-xe^{-x}]_0^\infty - \int_0^\infty -e^{-x}\,dx\\ &= 0 + [-e^{-x}]_0^\infty\\ &= 1.\end{aligned}$$

To finish, notice that $\log 2 < 1$.

27. A pdf f is called unimodal with mode a if $y \le x \le a$ implies $f(y) \le f(x) \le f(a)$ and $a \le x \le y$ implies $f(a) \ge f(x) \ge f(y)$.

a. *Give an example of a unimodal pdf for which the mode is unique.*

Any normal distribution would do.

b. *Give an example of a unimodal pdf for which the mode is not unique.*

You could take any uniform distribution here.

c. *Show that if f is symmetric about a and unimodal, then a is a mode.*

Assume that a were not a mode, so that there exists some $\epsilon \neq 0$ such that $f(a+\epsilon) > f(a)$. But then $f(a-\epsilon) > f(a)$ as well, which contradicts unimodality.

d. *Show that $f(x) = \mathbb{1}_{x \geq 0}\, e^{-x}$ is unimodal.*

Indeed, it has mode 0, since $0 \leq 1$ and f is strictly decreasing for $x > 0$.

28. *Let μ_n denote the nth central moment of X. Then recall*

$$\begin{aligned} \text{the skewness} \quad \alpha_3 &= \frac{\mu_3}{(\mu_2)^{3/2}} \\ \text{the kurtosis} \quad \alpha_4 &= \frac{\mu_4}{\mu_2^2}. \end{aligned}$$

a. *Let the pdf f be symmetric about a. Then $\alpha_3 = 0$.*

It suffices to show that $\mu_3 = 0$. Indeed, let $Y = X - a$ so that Y has pdf $f(y+a)$ and $EY^3 = \mu_3$. But then

$$EY^3 = \int_{-\infty}^{\infty} y^3 f(y+a)\, dy \quad = \int_0^{\infty} (-y^3 f(-y+a) + y^3 f(y+a))\, dy = 0.$$

b. *Calculate α_3 for $f(x) = e^{-x}$, $x \geq 0$, which is skewed to the right.*

Above, we calculated $EX = 1$. Also, we have

$$\begin{aligned} EX^2 &= \int_0^{\infty} x^2 e^{-x}\, dx \\ &= \left[-x^2 e^{-x}\right]_0^{\infty} - \int_0^{\infty} -2xe^{-x}\, dx \\ &= 0 + 2EX \\ &= 2. \end{aligned}$$

Then the variance is

$$\mu_2 = \operatorname{var} X = E[X^2] - E[X]^2 = 1.$$

Finally,

$$\begin{aligned} E[(X-1)^3] &= \int_0^{\infty} (x-1)^3 e^{-x}\, dx \\ &= [-(x-1)^3 e^{-x}]_0^{\infty} - \int_0^{\infty} -3(x-1)^2 e^{-x}\, dx \\ &= -1 + 3 \operatorname{var} X = 2. \end{aligned}$$

Thus

$$\alpha_3 = 2.$$

c. *Calculate α_4 for each of the following pdfs.*

$N(0,1)$: Here $\mu_1 = 0$ and $\mu_2 = 1$. Then $\alpha_4 = \mu_4$. Later, if we let $u = x^2$, $du = 2x\,dx$, and

$$\begin{aligned}
\mu_4 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 e^{-x^2/2}\,dx \\
&= \sqrt{\frac{2}{\pi}} \int_0^{\infty} x^4 e^{-x^2/2}\,dx \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} u^{3/2} e^{-u/2}\,du \\
&= \frac{1}{\sqrt{2\pi}} \Big(\big[-2u^{3/2} e^{-u/2}\big]_0^{\infty} - \int_0^{\infty} -2\sqrt{u} e^{-u/2}\,du \Big) \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} 2\frac{3}{2} \sqrt{u} e^{-u/2}\,du \\
&= \frac{6}{\sqrt{2\pi}} \int_0^{\infty} x^2 e^{-x^2/2}\,du \\
&= 3 \operatorname{var} X = 3.
\end{aligned}$$

uniform$(-1,1)$: Here, $\mu_1 = 0$, and

$$\mu_2 = \int_{-1}^{1} \frac{x^2}{2}\,dx = \left[\frac{1}{6}x^3\right]_{-1}^{1} = \frac{1}{3}.$$

Finally,

$$\mu_4 = \int_{-1}^{1} \frac{x^4}{2}\,dx = \left[\frac{1}{10}x^5\right]_{-1}^{1} = \frac{1}{5},$$

so that $\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{9}{5}$.

double-exponential: here, $\mu_1 = 0$, and

$$\begin{aligned}
\mu_2 &= \int_{-\infty}^{\infty} \frac{1}{2} x^2 e^{-|x|}\,dx \\
&= \int_0^{\infty} x^2 e^{-x}\,dx \\
&= 2,
\end{aligned}$$

where we have calculated the last integral above. Sadly, we didn't calculate the 4th central moment, but it's just

$$\begin{aligned}
\mu_4 &= \int_0^{\infty} x^4 e^{-x}\,dx \\
&= \big[-e^{-x} x^4\big]_0^{\infty} - \int_0^{\infty} -e^{-x}\, 4x^3\,dx \quad = 0 + 4\Big(\big[-e^{-x} x^3\big]_0^{\infty} - \int_0^{\infty} -e^{-x}\, 3x^2\,dx \Big) \\
&= 12 \int_0^{\infty} x^2 e^{-x}\,dx \\
&= 24 \quad (= 4!).
\end{aligned}$$

Thus the kurtosis is $\alpha_4 = 6$.

29. *Factorial moments. I'm just going to do Poisson because these binomial coefficients are killing me.*

    a. *Calculate the factorial moment $E[X(X-1)]$ for the Poisson.*

    Then $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$, so

$$\begin{aligned} E[X(X-1)] &= e^{-\lambda} \sum_{x=0}^{\infty} \lambda^x \frac{x(x-1)}{x!} \\ &= \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} \\ &= \lambda^2. \end{aligned}$$

    b. *Use that to find the variance.*

    Well, $E[X(X-1)] = EX^2 - EX$, and we know that the Poisson has expectation λ by a similar calculation. Thus $EX^2 = \lambda^2 + \lambda$. Finally,

$$\operatorname{var} X = EX^2 - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

30. *Find the mgf for the following pdfs.*

    a. $f(x) = \frac{1}{c}$, $0 < x < c$.

    Here,

$$\begin{aligned} Ee^{tX} &= \int_0^c \frac{e^{tx}}{c}\, dx \\ &= \left[\frac{e^{tx}}{tc}\right]_0^c \\ &= \frac{e^{tc}-1}{tc}. \end{aligned}$$

    b. $f(x) = \frac{2x}{c^2}$, $0 < x < c$.

$$\begin{aligned} Ee^{tX} &= \frac{2}{c^2}\int_0^c x\, e^{tx}\, dx \\ &= \frac{2}{c^2}\left(\left[\frac{1}{t}xe^{tx}\right]_0^c - \int_0^c e^{tx}\, dx\right) \\ &= \frac{2e^{ct}}{ct} - \frac{2}{c^2}\left[\frac{1}{t}e^{tx}\right]_0^c \\ &= \frac{2e^{ct}}{ct} - \frac{2}{c^2 t}(e^{ct}-1). \end{aligned}$$

    c. $f(x) = \frac{1}{2\beta}e^{-|x-\alpha|/\beta}$, $-\infty < x < \infty$, $-\infty < \alpha < \infty$, $\beta > 0$.

To simplify a little, let $Y = X - \alpha$, so that Y has pdf $g(y) = \frac{1}{2\beta}e^{-|y|/\beta}$. Then we can recover

$$\begin{aligned} M_X(t) &= e^{-\alpha t} E e^{tY} \\ &= \frac{e^{-\alpha t}}{2\beta} \int_{-\infty}^{\infty} e^{ty}\, e^{-|y|/\beta}\, dy. \end{aligned}$$

Let's break this into two integrals,

$$\begin{aligned} \int_{-\infty}^{0} e^{(t+\frac{1}{\beta})y}\, dy &= \left[\frac{e^{(t+\frac{1}{\beta})y}}{t+\frac{1}{\beta}} \right]_{-\infty}^{0} = \frac{1}{t+\frac{1}{\beta}}, \\ \int_{0}^{\infty} e^{(t-\frac{1}{\beta})y}\, dy &= \left[\frac{e^{(t-\frac{1}{\beta})y}}{t-\frac{1}{\beta}} \right]_{0}^{\infty} = \frac{1}{t-\frac{1}{\beta}}, \end{aligned}$$

where we only have convergence for $t < \frac{1}{\beta}$. In the end,

$$M_X(t) = \frac{e^{-\alpha t}}{2\beta} \left[\frac{1}{t+\frac{1}{\beta}} + \frac{1}{t-\frac{1}{\beta}} \right] = \frac{e^{-\alpha t}}{\beta} \frac{1}{t^2 - \beta^{-2}}.$$

31. *Does a distribution exist for which $M_X(t) = t/(1-t)$, $|t| < 1$?*

    No, because what random variable satisfies $Ee^{0X} = 0$?

32. *Let $S(t) = \log M_X(t)$. Then $S'(0) = EX$ and $S''(0) = \operatorname{var} X$.*

    Indeed,

    $$\frac{d}{dt} S(t)\bigg|_{t=0} = \frac{1}{M_X(t)} M_X'(t)\bigg|_{t=0} = 1 \cdot M_X'(0) = EX,$$

    and

    $$\begin{aligned} \frac{d^2}{dt^2} S(t)\bigg|_{t=0} &= \frac{d}{dt} \frac{M_X'(t)}{M_X(t)}\bigg|_{t=0} \\ &= \frac{M_X(0) M_X''(0) - M_X'(0)^2}{M_X(0)^2} \\ &= E[X^2] - E[X]^2. \end{aligned}$$

33. *Verify the following mgfs and use them to calculate EX and* $\operatorname{var} X$.

    a. Poisson(λ).

       Here,

       $$\begin{aligned} M_X(t) &= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} \\ &= e^{e^t\lambda - \lambda} \sum_{x=0}^{\infty} \frac{e^{-e^t\lambda}(e^t\lambda)^x}{x!} \\ &= e^{\lambda(e^t - 1)}. \end{aligned}$$

By the last exercise, we can consider $S_X(t) = \log M_X(t) = \lambda(e^t - 1)$. Then, interestingly, $S_X^{(n)}(t) = \lambda e^t$ for all $n > 0$, so that

$$EX = S'_X(0) = \lambda,$$

and

$$\operatorname{var} X = S''_X(0) = \lambda.$$

b. Geometric with x failures.

Using the geometric sum formula $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$, we have

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{\infty} e^{tx} p(1-p)^x \\ &= p \sum_{x=0}^{\infty} (e^t(1-p))^x \\ &= \frac{p}{1-e^t(1-p)}. \end{aligned}$$

Then, letting $q = 1-p$ to clean it up,

$$\begin{aligned} EX &= M'_X(t)\Big|_{t=0} = \frac{pqe^t}{(1-e^tq)^2}\Big|_{t=0} = \frac{1-p}{p}, \\ \operatorname{var} X &= M''_X(t)\Big|_{t=0} \\ &= p(1-p)\ (1-qe^t)^2e^t + 2qe^{2t}(1-qe^t)(1-qe^t)^4\Big|_{t=0} \\ &= p(1-p)\frac{p^2 + 2p(1-p)}{p^4} \\ &= \frac{1-p}{p^2}. \end{aligned}$$

c. Gaussian $N(\mu, \sigma^2)$.

Let's start with $N(0,1)$ and use the affine formula to finish. Completing the square in the exponent,

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x^2+2tx)/2}\,dx \\ &= e^{-t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x+t)^2/2}\,dx \\ &= e^{-t^2/2}. \end{aligned}$$

Then

$$M_{\sigma X+\mu} = e^{\mu t} M_X(\sigma t) = e^{\mu t + \sigma^2 t^2/2}.$$

34. *A distribution cannot be determined by a finite collection of moments. Let $X \sim N(0,1)$ and Y be the discrete random variable with $P(Y = \sqrt{3}) = P(Y = -\sqrt{3}) = \frac{1}{6}$ and $P(Y = 0) = \frac{2}{3}$. Then $EX^r = EY^r$ for $r = 1,2,3,4,5$.*

    Indeed, clearly the odd moments are all 0 since these distributions are symmetric. So, it suffices to compute the moments for $r = 2,4$. As we've computed previously, $EX^2 = 1$ and $EX^4 = 3$.

    For Y, we have $\frac{2\cdot 3}{6} = 1$ and $\frac{2\cdot 9}{6} = 3$.

35. *I'm including (a) of this problem below with 36.*

36. *Let X be standard lognormal, so that it has pdf*

$$f(x) = \frac{1}{\sqrt{2\pi x}} e^{-\log(x)^2/2}.$$

Show that $EX^r = e^{r^2/2}$, so that all moments exists, and show that the mgf does not exist.

Indeed, make the change of variables $y = \log x$ (so that $x^r = e^{ry}$) and $dy = \frac{dx}{x}$. Then,

$$\begin{aligned} EX^r &= \int_0^\infty \frac{x^r}{\sqrt{2\pi x}} e^{-\log(x)^2/2}\, dx \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2+ry}\, dy \\ &= e^{-r^2/2} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-(y^2-2ry+r^2)/2}\, dy \\ &= e^{-r^2/2}, \end{aligned}$$

where we recognized the $N(r,1)$ pdf in there. For the mgf, it suffices to show that the integral

$$\int_0^\infty \frac{e^{tx}}{x} e^{-\log(x)^2/2}\, dx$$

does not converge for any $t \neq 0$, since the mgf needs to exist in a neighborhood of 0. For this, it suffices to show that the integrand does not converge to 0. Indeed,

$$\begin{aligned} \lim_{x\to\infty} \frac{e^{tx}}{xe^{\log(x)^2/2}} &= \lim_{x\to\infty} \frac{te^{tx}}{e^{\log(x)^2/2} + \frac{1}{2}e^{\log(x)^2/2}} \\ &= 2 \lim_{x\to\infty} \frac{te^{tx}}{e^{-\log(x)^2/2}} \\ &= t\frac{2}{3} \lim_{x\to\infty} e^{tx-\log(x)^2/2} \\ &= t\frac{2}{3} \exp\left(\lim tx - \log(x)^2/2\right) = \infty. \end{aligned}$$

38. *Let X have the negative binomial distribution with pmf*

$$f(x) = \binom{r+x-1}{x} p^r (1-p)^x,$$

where $x = 0, 1, 2, \ldots$, $0 < p < 1$, and r is the number of successful outcomes.

a. *Calculate $M_X(t)$.*

$$\begin{aligned}
Ee^{tX} &= \sum_{x=0}^{\infty} e^{tx} \binom{r+x-1}{p}^r (1-p)^x \\
&= \frac{p^r}{(1-e^t+pe^t)^r} \sum_{x=0}^{\infty} \binom{r+x-1}{(} 1-e^t+pe^t)^r (e^t(1-p))^x \\
&= \frac{p^r}{(1-e^t+pe^t)^r}.
\end{aligned}$$

b. *Let $Y = 2pX$. Then the mgf of Y converges to*

$$\left(\frac{1}{1-2t}\right)^r$$

as $p \to 0$, which is the mgf of the χ^2_{2r} distribution.

Indeed, if $Y = 2pX$, then

$$M_Y(t) = M_X(2pt) = \left(\frac{p}{1-e^{2pt}+pe^{2pt}}\right)^r.$$

In the limit as $p \to 0$, we can pass inside the rth power and apply L'Hôpital to arrive at

$$\lim_{p\to\infty} \frac{p}{1-e^{2pt}+pe^{2pt}} = \lim_{p\to\infty} \frac{1}{-2te^{2pt}+e^{2pt}+2p^2e^{2pt}} = \frac{1}{1-2t},$$

which shows the result after restoring the rth power.

Chapter 3

Common Families of Distributions

1. *Let $X \sim$ Uniform$(\{a, \dots, b\})$. Find EX and* $\operatorname{var} X$.

   An old classic tells us that $EX = \frac{a+b}{2}$.

   For the variance, that result is not changed by shifting $Y = X - a + 1$ to recover the uniform distribution on $\{1, \dots, b-a\}$. But that distribution has second moment

   $$EY^2 = \frac{1}{b-a} \sum_{i=1}^{b-a} i^2 = \frac{(b-a+1)(2b-2a+1)}{6},$$

   for a final variance of

   $$\begin{aligned} \operatorname{var} X &= E[Y^2] - E[Y]^2 \\ &= \frac{(b-a+1)(2b-2a+1)}{6} - \frac{b-a+1}{2} \\ &= \frac{(b-a+1)(2b-2a-2)}{6} \\ &= \frac{(b-a+1)^2}{3}. \end{aligned}$$

3. *Let the event that a car passes in the ith second be modeled as the Bernoulli random variable $X_i \sim_{iid}$ Bernoulli(p). A pedestrian can only cross if no car is to pass in the next three seconds. Find the probability that the pedestrian has to wait 4 seconds before starting to cross.*

   The pedestrian has to wait 4 seconds before starting to cross if the following conditions obtain: $X_5 = X_6 = X_7 = 0$, and that there is no run of 3 0s in X_1, X_2, X_3, X_4. By independence, we just need to find $\frac{1}{8}$ times the probability of the second condition.

The only way the second condition can fail is if one of the following three events hold: $X_1 = X_2 = X_3 = X_4 = 0$, $X_1 = 1$ and $X_2 = X_3 = X_4 = 0$, or $X_1 = X_2 = X_3 = 0$ and $X_4 = 1$. Each of these have probability $\frac{1}{16}$, so the overall probability of interest is

$$\frac{1}{8}\left(1 - \frac{3}{16}\right) = \frac{13}{128}.$$

4. *A man with n keys tries to open a door by picking keys at random. Exactly one key will open the door. Find the mean number of trials if*

   a. *unsuccsessful keys are not eliminated from further selections.*

      Then we have a geometric random variable with success on the kth trial and $p = \frac{1}{n}$. This has mean $\frac{1}{p} = n$.

   b. *unsuccessful keys are eliminated.*

      To compute the expectation, we can do some telescoping:

$$\begin{aligned} EX &= \sum_{x=1}^{n} \frac{x}{n-x} \prod_{i=1}^{x-1} \frac{n-i}{n-i+1} \\ &= \sum_{x=1}^{n} \frac{x}{n-x} \frac{n-x}{n-x+1} \frac{n-x+1}{n-x+2} \cdots \frac{n-1}{n} \\ &= \sum_{x=1}^{n} \frac{x}{n} \\ &= \frac{n+1}{2}. \end{aligned}$$

5. *A standard drug is known to be effective in 80% of cases in which it is used. A new drug is tested on 100 patients and found to be effective in 85. What is the probability that the new drug is superior?*

   Equivalently, what is the probability that the old drug would perform at least as well in a similar test? Well, if $X \sim \text{Binomial}(n = 100, p = 0.8)$, then this is $P(X \geq 85)$.

   This is hard to calculate, although you could calculate something similar by taking $X \approx Y \sim N(80, 16)$. If $Z = \frac{Y-80}{4}$, $Y \geq 85$ implies $Z \geq 1.25$. They say 1.125 in the solutions, but that's not right.

6. *A large number of insects are expected to be attracted to a certain variety of rose plant. A commerical insecticide is advertised as 99% effective. Suppose 2000 insects infest a rose garden where the insecticide has been applied, and let X = number of surviving insects.*

   a. *What is a reasonable model?*

      I guess we should choose $X \sim \text{Binomial}(n = 2000, p = 0.01)$.

b. *Write an expression for the probability that fewer than 100 insects survive.*

$$P(X < 100) = \sum_{i=0}^{9} 9\binom{2000}{i} 0.01^i 0.99^{2000-i}.$$

c. *Evaluate an approximation to that probability.*

If we take $X \approx Y \sim N(20, 19.8)$, we want to find the probability that $Y \leq 100$, approximated by the probability that

$$Z = \frac{X - 20}{\sqrt{19.8}} \leq \frac{100 - 20}{\sqrt{19.8}} \approx 17.97.$$

This is clearly going to be extremely close to 1.

7. *Let the number of chocolate chips in a cookie have a Poisson distribution. We want the probability that a random cookie has at least two chips to be larger than 0.99. Find the smallest value of the mean of the distrubution to ensure this.*

Indeed, let $X \sim \text{Poisson}(\lambda)$, and recall $EX = \lambda$. Then

$$P(X \leq 1) = \sum_{x=0}^{1} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}(1 + \lambda).$$

Then the smallest possible λ will satisfy

$$1 + \lambda = 0.01e^{\lambda}$$

which works for $\lambda \approx 6.64$, using a very fancy calculator.

13. *Let* $X \sim \text{Poisson}(\lambda)$. *What are the pmf, mean, and variance of the 0-truncated version?*

The bulk is noticing that $P(X > 0) = 1 - e^{-\lambda}$. That becomes the denominator...

15. *Let X be a negative binomial*(r, p) *random variable. Then as* $r \to \infty$ *and* $p \to 1$, *if* $r(1 - p) \to \lambda$, *then* $M_X \to M_Y$ *where* $Y \sim \text{Poisson}(\lambda)$.

Indeed, the Poisson has moment generating function

$$\begin{aligned} M_Y(t) &= Ee^{tY} \\ &= \sum_{y=0}^{\infty} \frac{e^{ty}e^{-\lambda}\lambda^y}{y!} \\ &= \frac{e^{-\lambda}}{e^{-e^t\lambda}} \sum_{y=0}^{\infty} \frac{e^{-e^t\lambda}(e^t\lambda)^y}{y!} \\ &= e^{e^t\lambda - \lambda}. \end{aligned}$$

The negative binomial has mgf

$$\begin{aligned}
M_X(t) &= Ee^{tX} \\
&= p^r \sum_{x=0}^{\infty} \binom{r+x-1}{r-1} e^{tx} q^x \\
&= \left(\frac{p}{1-e^t q}\right)^r \sum_{x=0}^{\infty} \binom{r+x-1}{r-1} (1-e^t q)^r (e^t q)^x \\
&= \left(\frac{p}{1-(1-p)e^t}\right)^r.
\end{aligned}$$

I assume there's some way to massage this into a form that looks sort of like $(1+\frac{\lambda(e^t-1)}{r})^r$, at least asymptotically.

Chapter 4

Multiple Random Variables

1. *Let (X, Y) be uniformly distributed on the square $[-1, 1] \times [-1, 1]$. Determine the probabilities of the following events.*

   a. $X^2 + Y^2 < 1$.

      The probability of landing in the unit circle is $\frac{\pi}{4}$.

   b. $2X - Y > 0$.

      Any line through the origin cuts the square in half.

   c. $|X + Y| < 2$.

      All points except 4 in the square are contained by this larger diamond.

4. *Consider the pdf*

$$f(x, y) = C(x + 2y) \quad \text{if} \quad 0 < y < 1, 0 < x < 2.$$

   a. *Find C.*

      We have

$$\begin{aligned}
\int_0^1 \int_0^2 x + 2y \, dxdy &= \int_0^1 \left[\frac{x^2}{2} + 2yx \right]_{x=0}^2 dy \\
&= \int_0^1 2 + 4y \, dy \\
&= \left[2y + 2y^2 \right]_0^1 \\
&= 4,
\end{aligned}$$

      so that $C = \frac{1}{4}$.

   b. *Find the marginal distribution of X.*

Integrating out Y,

$$\begin{aligned} f_X(x) &= \frac{1}{4}\int_0^1 x + 2y\,dy \\ &= \frac{1}{4}\left[xy + y^2\right]_{y=0}^1 \\ &= \frac{x+1}{4}. \end{aligned}$$

c. *Find the join cdf.*

Since X and Y start at 0, there is no integration constant, and

$$F(x,y) = \frac{1}{4}\left(\frac{x^2y}{2} + xy^2\right).$$

d. *Find the pdf of the random variable $Z = 9/(X+1)^2$.*

Write $h(x) = 9/(x+1)^2$. As X varies from 0 to 2, we see that the support of Z is $[1,9]$. h is differentiable and invertible there, so that

$$f_Z(z) = f_X(h^{-1}(z))\left|\frac{d}{dz}h^{-1}(z)\right|.$$

Since

$$h^{-1}(z) = 3z^{-\frac{1}{2}} - 1 \implies f_X(h^{-1}(z)) = \frac{3}{4\sqrt{z}}$$

and

$$\frac{d}{dz}h^{-1}(z) = -\frac{3}{2z^{3/2}},$$

we have

$$f_Z(z) = \frac{9}{8z^2}.$$

5. a. *Let X and Y have joint pdf $f(x,y) = x + y$ on $[0,1]^2$. Find $P(X > \sqrt{Y})$.*

$$\begin{aligned} P(X > \sqrt{Y}) &= \int_0^1 \int_{\sqrt{y}}^1 x + y\,dx\,dy \\ &= \int_0^1 \left[\frac{1}{2}x^2 + xy\right]_{x=\sqrt{y}}^1 dy \\ &= \int_0^1 \frac{1}{2} + \frac{y}{2} - y^{3/2}\,dy \\ &= \left[\frac{y}{2} + \frac{y^2}{4} - \frac{2}{5}y^{5/2}\right]_0^1 \\ &= \frac{1}{2} + \frac{1}{4} - \frac{2}{5} = \frac{7}{20}. \end{aligned}$$

b. *Now let them be jointly distributed with pdf $f(x,y) = 2x$. What is $P(X^2 < Y < X)$?*

$$
\begin{aligned}
P(X^2 < Y < X) &= \int_0^1 \int_{x^2}^{x} 2x \, dy \, dx \\
&= \int_0^1 [2xy]_{y=x^2}^{x} \, dx \\
&= \int_0^1 2x^2 - 2x^3 \, dx \\
&= \left[\frac{2}{3}x^3 - \frac{1}{2}x^4 \right]_0^1 \\
&= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.
\end{aligned}
$$

6. *A and B agree to meet at a certain place between 1pm and 2pm. Say they arrive independently and uniformly during the hour. Find the distribution of the length of time that A waits for B, counted as 0 if B arrives before A.*

   Let X and Y be the arrival times of A and B, expressed in fractions of hours so that they are both uniform on $[0,1]$ with joint density $f(x,y) = 1$. Define A's waiting time $Z = \max(Y - X, 0)$. Note that Z takes on the value 0 with probability $\frac{1}{2}$ since they are equally likely to arrive first.

   Drawing squares and triangles, you can see that $P(Z > z) = \frac{1}{2}(1 - z^2)$. So,

$$
P(Z \leq z) = \begin{cases} \frac{1}{2} & z = 0 \\ \frac{1}{2}(1+z)^2 & z > 0. \end{cases}
$$

7. *A woman leaves for work between 8 and 8:30, and takes 40 to 50 minutes to get there. What's the probability she arrives before 9?*

   Let X be uniform$(0, 30)$ and Y uniform $(40, 50)$. Then we want $P(X + Y < 60)$. This is just...

$$
\begin{aligned}
P(X + Y < 60) &= \int_4 0^5 0 \int_0^{60-y} \frac{1}{300} \, dx \, dy \\
&= \int_4 0^5 0 \frac{60 - y}{300} \, dy \\
&= \left[\frac{120y - y^2}{600} \right]_4 0^5 0 \\
&= \frac{1}{600}(3500 - 3200) \\
&= \frac{1}{2}.
\end{aligned}
$$

   Thinking about it more, we can notice that $E[X + Y]$ tells us she gets there at 9 on average, and we can notice that the distribution is symmetric about the mean to arrive at the same answer.

9. *Let X,Y such that $F_{X,Y}(x,y) = F_X(x)F_Y(y)$. Then for any pair of intervals (a,b), (c,d),*

$$P(a \le X \le b, c \le Y \le d) = P(a \le X \le b)P(c \le Y \le d).$$

Indeed, by inclusion-exclusion,

$$\begin{aligned} P(a \le X \le b, c \le Y \le d) &= P(X \le b, Y \le d) - P(X \le a) - P(Y \le c) + P(X \le a, Y \le c) \\ &= F_{X,Y}(b,d) - F_{X,Y}(a,c) - F_{X,Y}(b,c) + F_{X,Y}(a,c). \end{aligned}$$

and from the other direction,

$$\begin{aligned} P(a \le X \le b)P(c \le Y \le d) &= (F_X(b) - F_X(a))(F_Y(d) - F_Y(c)) \\ &= F_X(b)F_Y(d) + F_X(a)F_Y(c) - F_X(b)F_Y(c) - F_X(a)F_Y(d) \\ &= F_{X,Y}(b,d) + F_{X,Y}(a,c) - F_{X,Y}(b,c) - F_{X,Y}(a,d). \end{aligned}$$

10. *X, Y have a joint distribution that I can't figure out how to get Pandoc to render. Oh wait, here we go.*

			X	
		1	2	3
	2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
Y	3	$\frac{1}{6}$	0	$\frac{1}{6}$
	4	0	$\frac{1}{3}$	0

a. *Show that X and Y are dependent.*

Indeed,

$$P(X = 2)P(Y = 3) = \frac{1}{3}\frac{1}{3} \neq 0 = P(X = 2, Y = 3).$$

b. *Give a distribution for independent random variables U and V with the same marginals.*

Indeed, the marginals are

y	$P(Y = y)$	x	$P(X = x)$
2	$\frac{1}{3}$	1	$\frac{1}{4}$
3	$\frac{1}{3}$	2	$\frac{1}{2}$
4	$\frac{1}{3}$	3	$\frac{1}{4}$

By independence, we can directly take products to get the diagonal.

			X	
		1	2	3
	2	$\frac{1}{12}$	a	b
Y	3	c	$\frac{1}{6}$	d
	4	e	f	$\frac{1}{12}$

The rest of the elements satisfy the system of equations

$$\begin{aligned}
\frac{1}{4} &= a + b \\
\frac{1}{6} &= c + d \\
\frac{1}{4} &= e + f \\
\frac{1}{6} &= c + e \\
\frac{1}{3} &= a + f \\
\frac{1}{6} &= b + d.
\end{aligned}$$

This is not full-rank, there are infinitely many solutions. You can pick one, that's fine with me.

11. *Let U = number of trials needed to get first head, V = number of trials needed to get two heads. Are U and V independent?*

    No, they're not. Consider that $P(U = 4, V = 2) = 0$ while $P(U = 4)$ and $P(V = 2)$ are both > 0.

12. *If a stick is broken at random into three pieces, what is the probability that the pieces can be put together in a triangle?*

    I really wonder how this depends on the breaking procedure. For instance, we could model this as: let X, Y be two uniform $(0, 1)$ breaking points. Then the lengths of the segments are $A = min(X, Y)$, $B = \max(X, Y) - \min(X, Y)$ and $C = 1 - \max(X, Y)$.

    I think an easier (but clearly equivalent) model is to let (A, B, C) be uniformly chosen on the set $\{(a, b, c) : 0 < a, b, c < 1, a + b + c = 1\}$. To do so, let (A, B) be uniformly distributed on the triangle $a < b$, $0 < a, b < 1$ (with density $1/2$), and pick $C = 1 - A - B$.

    Then we want the probability that these numbers satisfy the triangle inequality:

    $$\begin{aligned}
    P(\text{satisfy triangle inequality}) &= P(A < B + C, A > B, A > C) + P(B < A + C, B > A, B > C \\
    &= 3P(C < A + B \mid C > A, C > B) \\
    &= 3 \int_0^1 \int_a^1 \mathbb{1}_{1-a-b>a+b} \, \frac{1}{2} dbda.
    \end{aligned}$$

    Thinking about this integral geometrically, one identifies a tiny triangle with area $\frac{1}{16}$. Multiplied by the 3 and the measure $\frac{1}{2}$, we get the answer $\frac{3}{32}$. Huh. This is not the answer they got. But they really are looking at a different set. Weird...

13. *Let X and Y be random variables with finite means. Then*

    $$\min_g E[(Y - g(X))^2] = E[(Y - E[Y \mid X])^2].$$

Indeed, for all g,

$$\begin{aligned} E[(Y-g(X))^2] &= E[(Y-E[Y \mid X]+E[Y \mid X]-g(X))^2] \\ &= E[(Y-E[Y \mid X])^2]+E[(E[Y \mid X]-g(X))^2]+2E[(Y-E[Y \mid X])(E[Y \end{aligned}$$

If we can show that the last term is 0, we'll be done. Indeed, we can add an $E[\cdot \mid X]$ without changing this term's value, to see that

$$\begin{aligned} E[(Y-E[Y \mid X])(E[Y \mid X]-g(X))] &= E[E[(Y-E[Y \mid X]) \mid X]] \\ &= E[(E[Y \mid X]-g(X))E[Y-E[Y \mid X] \mid X]] \\ &= E[(E[Y \mid X]-g(X))(E[Y \mid X]-E[Y \mid X])] \\ &= 0. \end{aligned}$$

14. *Let $X, Y \sim_{iid} N(0,1)$.*

   a. *Find $P(X^2+Y^2<1)$.*

   Recall that we have the joint density

$$f(x,y) = \frac{1}{\sqrt{2\pi}} e^{-(x^2+y^2)/2}.$$

   So, we can go to polar coordinates to get

$$\begin{aligned} P(X^2+Y^2<1) &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2}\, r\, dr\, d\theta \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \int_0^{\infty} e^{-s}\, ds\, d\theta \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} [-0+1]\, d\theta \\ &= \sqrt{2\pi}. \end{aligned}$$

   where we made the change of variables $s = r^2/2$, $ds = rdr$.

   b. *Verify that $X^2 \sim \chi_1^2$ and find $P(X^2<1)$.*

   Indeed,

$$\begin{aligned} \frac{d}{dx} P(X^2 \le x) &= \frac{d}{dx} P(-\sqrt{x} \le X \le \sqrt{x}) \\ &= \frac{d}{dx} \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \\ &= 2 \frac{1}{2\sqrt{x}\sqrt{2\pi}} e^{-x/2}. \end{aligned}$$

   Recall that the χ_p^2 has pdf

$$f(\chi \mid p) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} e^{-x/2},$$

which is great since $\Gamma(1/2) = \sqrt{\pi}$.

Directly integrating this will turn into some Γ function, or just a Gaussian probability integral. Not sure what to do here...

15. *Let $X \sim \text{Poisson}(\theta)$, $Y \sim \text{Poisson}(\lambda)$, independent. Then $X \mid X + Y$ is binomial with success probability $\theta/(\theta + \lambda)$.*

    Indeed, by independence, and since $X + Y \sim \text{Poisson}(\theta + \lambda)$,

$$\begin{aligned}
P(X = x \mid X + Y = n) &= \frac{P(X = x, Y = n - x)}{P(X + Y = n)} \\
&= \frac{e^{-\theta}\theta^x}{x!} \frac{e^{-\lambda}\lambda^{n-x}}{(n-x)!} \frac{n!}{e^{-(\theta+\lambda)}(\theta+\lambda)^n} \\
&= \frac{e^{-\theta}e^{-\lambda}}{e^{-(\theta+\lambda)}} \frac{n!}{x!(n-x)!} \frac{\theta^x}{(\theta+\lambda)^x} \frac{\lambda^{n-x}}{(\theta+\lambda)^n - x} \\
&= \binom{n}{x} \left(\frac{\theta}{\theta+\lambda}\right)^x \left(\frac{\lambda}{\theta+\lambda}\right)^{n-x}.
\end{aligned}$$

    To finish, note that $1 - \theta/(\theta + \lambda) = \lambda/(\theta + \lambda)$.