Statistics 610 Assignment 2
due Friday, 12 September 2025

Note: you may use a computer or calculator to do the numerical computations and for the plots.

1. Chapter 1 Exercise 34.

2. (a) Prove this alternative to Bayes' rule:

$$\log\left(\frac{\mathsf{P}(A|B)}{\mathsf{P}(A^c|B)}\right) = \log\left(\frac{\mathsf{P}(A)}{\mathsf{P}(A^c)}\right) + \log\left(\frac{\mathsf{P}(B|A)}{\mathsf{P}(B|A^c)}\right).$$

   This expression is useful in genetics: conditional log odds of disease (A) given gene (B) = unconditional log odds of disease + log ratio of gene prevalence.

   (b) Suppose $B_1, \ldots, B_n$ are disjoint. Show that

$$\mathsf{P}\Big(B_j | \bigcup_{i=1}^{n} B_i\Big) = \frac{\mathsf{P}(B_j)}{\sum_{i=1}^{n} \mathsf{P}(B_i)} \quad \text{for each } j \in \{1, \ldots, n\}.$$

3. Yule's Paradox. A new drug is introduced to combat a disease. Let R be the event the drug is applied and I be the event the patient improves.

   (a) Suppose the following proportions are reported, based on an experimental sample of patients.

	R	R^c	
I	.412	.276	.688
I^c	.155	.157	.312
	.567	.433	1.000

   Find the chance the patient improves given that he takes the drug. Find the chance of improvement when he does not get the drug. What does this say about the efficacy of the drug?

   (b) What the researchers *did not* report is that patients can easily be categorized into two groups: those in early stages of the disease (E) and those not (E^c). The proportions then become as follows.

	$E \cap R$	$E \cap R^c$	$E^c \cap R$	$E^c \cap R^c$
I	.392	.216	.020	.060
I^c	.077	.039	.078	.118
	.469	.255	.098	.178

   Now compare the chance of improvement with and without the drug just for the patients in the early stages. Do the same for the patients in later stages. What do you think of the efficacy of the drug now? How should this experiment have been designed differently? ("Design" here means "number of patients chosen for each group in the study".)

4. Consider sampling with replacement from a population of size N in which M individuals would respond Yes (and $N - M$ would respond No).

   (a) Show that the first and second responses are independent. As the responses are binary, it suffices to show that events A_i = "i-th response is Yes" are independent, $i = 1, 2$.

   (b) Show that the n responses for a sample of size n are *mutually* independent. As in part (a), you can simply show that $A_1, \ldots, A_n$ are mutually independent.

continued next page

5. Consider rolling two fair dice.

   (a) Define events A = "first die is 6", B = "second die is 2", C = "total of the dice is 7". Are the events pairwise independent? Are they mutually independent?

   (b) Now consider A = "first die is even", B = "the two dice are the same", C = "total is 8, 9 or 10". Are the events pairwise independent? Are they mutually independent?

   (c) Suppose A, B, C are *pairwise* independent events, not necessarily mutually independent. Must A be independent either of $B \cap C$ or of $B \cup C$? Prove or show counterexamples.

6. Two shaved dice are such that the chance of a 1 or 6 is 1/5 each, and the chance of a 2, 3, 4 or 5 is 3/20 each.

   (a) Find the chance a total of 7 is obtained when the dice are rolled.

   (b) Now supposed the dice are repeatedly rolled independently for a total of 6 times. Let Y be the number of times that a total of 7 is obtained. Find the pmf and cdf of Y and plot both. (A line graph works best for a pmf: draw a line or narrow bar of height $f_Y(x)$ at value x. Also, be sure to explicitly indicate the cdf's value at each of the jump points with, say, a small darkened circle.)

7. Chapter 1 Exercise 51. Plot the pmf also.

8. Suppose $g(x) = 4\left(\frac{1}{3}\right)^x + \left(\frac{2}{3}\right)^x$ for $x = 1, 2, \ldots$, and $g(x) = 0$ otherwise.

   (a) Find $c > 0$ such that $f(x) = cg(x)$ is a valid pmf.

   (b) Find the corresponding cdf.

   (c) Consider this game: A fair coin is flipped. If a Head appears then a fair die is rolled until either a 1 or a 6 appears. If a Tail is flipped then the die is rolled until a 2, 3, 4 or 5 appears. Let X be the number of rolls and show that f (the function in part (a)) is its pmf.