Stat 110 Homework 5, Fall 2013

Due: Friday 10/11 at 1:10 pm, in the Stat 110 dropbox outside of SC 105. No late homework will be accepted. Please write your name, and staple your homework. Show your work and make sure to give clear, careful, convincing explanations.

Reading: Sections 4.7-4.8, 4.10, 5.1-5.4

1. Use Poisson approximations to investigate the following types of coincidences. The usual assumptions of the birthday problem apply, such as that there are 365 days in a year, with all days equally likely.

(a) How many people are needed to have a 50% chance that at least one of them has the same birthday as *you*?

(b) How many people are needed to have a 50% chance that there are two people who not only were born on the same day, but also were born at the same *hour* (e.g., two people born between 2 pm and 3 pm are considered to have been born at the same hour).

(c) Considering that only 1/24 of pairs of people born on the same day were born at the same hour, why isn't the answer to (b) approximately $24 \cdot 23$? Explain this intuitively, and give a simple approximation for the factor by which the number of people needed to obtain probability p of a birthday match needs to be scaled up to obtain probability p of a birthday-birthhour match.

(d) With 100 people, there is a 64% chance that there are 3 with the same birthday (according to R, using `pbirthday(100,classes=365,coincident=3)` to compute it). Provide two different Poisson approximations for this value, one based on creating an indicator r.v. for each triplet of people, and the other based on creating an indicator r.v. for each day of the year. Which is more accurate?

2. The number of people who visit the Leftorium store in a day is Pois(100). Suppose that 10% of customers are *sinister* (left-handed), and 90% are *dexterous* (right-handed). Half of the sinister customers make purchases, but only a third of the dexterous customers make purchases. The characteristics and behavior of people are independent, with probabilities as described in the previous two sentences. On a certain day, there are 42 people who arrive at the store but leave without making a purchase. Given this information, what is the conditional PMF of the number of customers on that day who make a purchase?

3. A circle with a random radius $R \sim \text{Unif}(0, 1)$ is generated. Let A be its area.

(a) Find the mean and variance of A, without first finding the CDF or PDF of A.

(b) Find the CDF and PDF of A.

4. (Exercise 5.14 in the book) Let $Z \sim \mathcal{N}(0,1)$ and $X = Z^2$. Then the distribution of X is called *Chi-Square with 1 degree of freedom*. This distribution appears in many statistical methods.

(a) Find a good numerical approximation to $P(1 \le X \le 4)$ using facts about the Normal distribution, without querying a calculator/computer/table about values of the Normal CDF.

(b) Let Φ and φ be the CDF and PDF of Z, respectively. Show that for any $t > 0$, $I(Z > t) \le (Z/t)I(Z > t)$. Using this and LOTUS, show that $\Phi(t) \ge 1 - \varphi(t)/t$.

5. Let $X_1, X_2, \dots$ be the annual rainfalls in Boston (measured in inches) in the years 2101, 2102, $\dots$, respectively. Assume that annual rainfalls are i.i.d. draws from a continuous distribution. A rainfall value is a *record high* if it is greater than those in all previous years (starting with 2101), and a *record low* if it is lower than those in all previous years.

(a) In the 22nd century (the years 2101 through 2200, inclusive), find the expected number of years that have either a record low or a record high rainfall.

(b) On average, in how many years in the 22nd century is there a record low followed in the next year by a record high?

(c) By definition, the year 2101 is a record high (and record low). Let N be the number of years required to get a new record high. Find $P(N > n)$ for all positive integers n, and find the PMF of N.