

MATH 20 – HOMEWORK 6 SOLUTIONS!

due Wednesday, August 9

Instructions: This assignment is due at the *beginning* of class. Staple your work together (do not just fold over the corner). Please write the questions in the correct order. If I cannot read your handwriting, you won't receive full credit.

You may use Wolfram Alpha to compute any necessary sums or integrals. If you have trouble with this, let me know.

If you're using facts about distributions to answer the questions, be very clear about which distribution you're using to model that problem and why that distribution is appropriate.

1. A school wishes to accept 2000 students for their freshman class, and they expect 20,000 applications. In order to make their admissions decisions very easy, the only criterion they will use is SAT score. So, their goal is to accept a student if and only if their SAT score is in the top 10%. However, because their computer system is so old, the applications only come in one at a time, and they must decide whether to accept or reject before moving on to the next application. Assuming that SAT scores are normally distributed with a mean of 1000 and a standard deviation of 200, how should they set the score threshold to end up with as close to 2000 students as possible? Give your answer first symbolically (in terms of a pdf, cdf, etc), then use a normal distribution table¹ to provide a numerical answer.

Solution: We want to find the SAT score x such that 90% of scores are below x and 10% of scores are above x . A symbolic answer can be expressed in many different ways. Here are a few (make sure you see why they're all equivalent):

- (a) Let $F(x)$ be the cumulative distribution function for the normal distribution with mean 1000 and standard deviation 200. Then, the x we want is the solution to the equation $F(x) = 0.9$.
- (b) Let $f(x)$ be the density function for the normal distribution with mean 1000 and standard deviation 200. Then, the x we want is the solution to

$$0.9 = F(x) = \int_{-\infty}^x \frac{1}{200\sqrt{2\pi}} e^{-\frac{(x-1000)^2}{80000}} dx.$$

- (c) Let $f(x)$ be the density function for the standard normal distribution with mean 0 and standard deviation 1. Then, the answer we want is $1000 + 200x$, where x is the solution to

$$0.9 = F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

The last characterization is the easiest to relate to the normal distribution table. We look through the table for the value closest to 0.90, and find that in a standard normal distribution, $P(X \leq 1.28) = 0.8997$ and $P(X \leq 1.29) = 0.9015$. We'll use the first value because its probability is closer to 0.9.

¹For example: <http://www.stat.ufl.edu/~athienit/Tables/Ztable.pdf>. The numbers in the leftmost column represent the number of standard deviations from the mean up to 1 decimal place, and the numbers along to top row then refine to the second decimal place. Please ask me if you have questions about this.

Hence, the cutoff should be placed 1.28 standard deviations above the mean. This is

$$1000 + 200(1.28) = 1256 \text{ points.}$$

2. The density function $f(x)$ of a continuous random variable X is given by

$$f(x) = \begin{cases} A + Bx^2, & \text{if } x \in [0, 2] \\ 0, & \text{otherwise} \end{cases}.$$

If $\mathbb{E}[X] = 1/2$, find A and B .

Solution: There are two unknowns, A and B , so we should look for two pieces of information. The first is that $\mathbb{E}[X] = 1/2$. The second is that $f(x)$ is actually a density, telling us that the integral of $f(x)$ is 1. So,

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x) dx \\ &= \int_0^2 (A + Bx^2) dx \\ &= \left[Ax + \frac{B}{3}x^3 \right]_0^2 \\ &= 2A + \frac{8}{3}B. \end{aligned}$$

Since, $\mathbb{E}[X] = 1/2$, we know

$$\begin{aligned} \frac{1}{2} &= \int_{-\infty}^{\infty} xf(x) dx \\ &= \int_0^2 x(A + Bx^2) dx \\ &= \left[\frac{A}{2}x^2 + \frac{B}{4}x^4 \right]_0^2 \\ &= 2A + 4B. \end{aligned}$$

Solve the simultaneous equations $2A + (8/3)B = 1$ and $2A + 4B = 1/2$, we find $A = 1$ and $B = -3/8$. Hence the density function is

$$f(x) = \begin{cases} 1 - \frac{8}{3}x^2, & \text{if } x \in [0, 2] \\ 0, & \text{otherwise} \end{cases}.$$

3. Since 1851, exactly 116 hurricanes have hit Florida (this includes the years 1851 and 2016, but not 2017—only direct hits by *hurricanes* are counted, not tropical storms). In 2005, Florida was hit by four hurricanes: Cindy, Dennis, Katrina, and Wilma. If the probability of hurricane strikes has remained the same since 1851, what is the probability of Florida being struck by four or more hurricanes in the same year?

Solution: This is a classic Poisson distribution. We've assumed that the probability of hurricane strikes has remained the same², hence our rate is 116 hurricanes per $2016 - 1851 + 1 = 166$ years. As the question asks about a year time frame, we have to adjust the rate:

$$\lambda = \frac{116/166 \text{ hurricanes}}{1 \text{ year}}.$$

²In reality, a bad assumption.

Now, the probability of four or more hurricanes in the same year is

$$1 - \sum_{k=0}^3 P(X = k) = 1 - \sum_{k=0}^3 \frac{\lambda^k e^{-\lambda}}{k!} \approx 0.005719 = 0.57\%.$$

In other words, this is a “1 in 200 years” type of event.

4. In the solutions manual to a Calculus textbook, there is about one faulty solution per fifty questions. In a book with ten chapters, each with one hundred questions, what is the probability that there are at least 15 faulty solutions in the whole book? Give your answer two ways: first with a binomial distribution, then with a Poisson approximation. Use Wolfram Alpha or some other tools to find both answers numerically, and compare them.

Solution: This is exactly a binomial distribution, and approximately a Poisson distribution.

A “success” is a faulty solution, hence $p = 1/50 = 0.02$. There are 1000 total problems, and so the probability that at least 15 are faulty is

$$P(X \geq 15) = \sum_{k=15}^{1000} \binom{1000}{k} \left(\frac{1}{50}\right)^k \left(\frac{49}{50}\right)^{1000-k} \approx 0.89747 = 89.7\%.$$

Using a Poisson approximation, the rate (which needs to have “per book” as its unit measurement) is

$$\lambda = \frac{20 \text{ faulty solutions}}{1 \text{ book}}.$$

Hence,

$$P(X \geq 15) = \sum_{k=15}^{1000} \frac{\lambda^k e^{-\lambda}}{k!} = 0.89513 = 89.5\%.$$
