Chris Piech
CS109

Section #4
October 24, 2018

Section #4 Solutions

1. **Are we due for an earthquake?**:

   a. What is the probability of no 8+ earthquakes in four years? Let X be the time until an earthquake. $X \sim \text{Exp}(\lambda = 0.002)$.

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) \\ &= 1 - F_X(4) \\ &= 1 - [1 - e^{-0.002 \cdot 4}] \\ &= e^{-0.008} \approx 0.992 \end{aligned}$$

   b. What is the probability of no 8+ earthquakes in the 113 years?

$$\begin{aligned} P(X \geq 113) &= 1 - P(X < 113) \\ &= 1 - F_X(113) \\ &= 1 - [1 - e^{-0.002 \cdot 113}] \\ &= e^{-0.226} \approx 0.798 \end{aligned}$$

   c. What is $P(X > 113 | X > 109)$?

$$\begin{aligned} P(X > 113 | X > 109) &= \frac{P(X > 113, X > 109)}{P(X > 109)} \\ &= \frac{P(X > 113)}{P(X > 109)} = \frac{1 - F_X(113)}{1 - F_X(109)} \\ &= \frac{e^{-0.002 \cdot 113}}{e^{-0.002 \cdot 109}} = e^{-0.008} \approx 0.992 \end{aligned}$$

   d. It turns out that exponentials are what we call a “memoryless distribution." If X is an exponential random variable, it holds that $P(X > s + t | X > t) = P(X > s)$.

2. **ReCaptcha**

   a. What the the probability density function of a robot clicking $X = x$ pixels from the left of the box and $Y = y$ pixels from the top of the box?

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{100]} & \text{if } 0 < x, y < 10 \\ 0 & \text{else} \end{cases}$$

b. Let $D \sim \text{Rayleigh}(\theta = 2)$ be the distance a human clicks.

$$\begin{aligned} P(D > 1.2) &= 1 - P(D < 1.2) = 1 - F_D(1.2) \\ &= 1 - [1 - e^{-1.2^2/2\cdot 2}] = e^{-1.2^2/4} \approx 0.698 \end{aligned}$$

c. We can start by expanding Bayes theorem for the new belief of a Robot

$$P(\text{Robot}|D = 2) = \frac{f(D = 2|\text{Robot})P(\text{Robot})}{f(D = 2)}$$

The two terms on the top are both ones that we can calculate from formulas that we have. The denominator is more problematic: it asks, what is the density of a click two pixels away if we don't know whether the user is a Robot or a Human. The answer is to use the law of total probability, just like in the past. Since all users are either humans or robots, $P(\text{Robot}|D = 2) + P(\text{Human}|D = 2) = 1$. As such:

$$\begin{aligned} P(\text{Robot}|D = 2) &= \frac{f(D = 2|\text{Robot})P(\text{Robot})}{f(D = 2|\text{Robot})P(\text{Robot}) + f(D = 2|\text{Human})P(\text{Human})} \\ &= \frac{\frac{1}{100} \cdot 0.2}{\frac{1}{100} \cdot 0.2 + \frac{2}{2} e^{-2^2/2\cdot 2} \cdot 0.8} \\ &= \frac{0.002}{0.002 + e^{-1} \cdot 0.8} \approx 0.006 \end{aligned}$$

3. It's Complicated

a. For each assignment to R, sum over all the values that S can take on that is consistent with that assignment.
Single = 0.44
In a Relationship = 0.47
It's Complicated = 0.09

b. Single = 0.125
In a Relationship = 0.875
It's Complicated = 0.00

c. Freshman = 0.35
Sophomore = 0.39
Junior = 0.42
Senior = 0.90
5+ = 0.60