SDS 321 Tips & Tricks

Evan Ott (Our friendly and awesome TA from 2016)

Here, I'm going to list out some things that you should use as sanity checks, along with some techniques to make your life easy.

- Probabilities are always between 0 and 1.
- Probability densities (pdfs of continuous variables) must be non-negative everywhere (0 or greater), and MUST integrate to 1. They should not be infinite, nor should they integrate to infinity.
- If a joint pdf (pdf with more than one variable) can be factored, the variables are independent:

$$\begin{aligned} f_{X,Y}(x,y) &= \begin{cases} xe^{-(x+y)} & x,y>0 \\ 0 & \text{otherwise} \end{cases} \\ &= \left(\begin{cases} xe^{-x} & x>0 \\ 0 & \text{otherwise} \end{cases} \right) \cdot \left(\begin{cases} e^{-y} & y>0 \\ 0 & \text{otherwise} \end{cases} \right) \\ &= f_X(x) f_Y(y) \end{aligned}$$

- We'll **never** ask you to do integration by parts; all integrals are going to be a function of a known pdf. For example:

$$\begin{aligned} f_{X,Y}(x,y) &= \begin{cases} xe^{-(x+y)} & x,y>0 \\ 0 & \text{otherwise} \end{cases} \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx \\ &= \int_0^{\infty} xe^{-(x+y)}dx = e^{-y}\int_0^{\infty} xe^{-x}dx \end{aligned}$$

Note: $e^{-x}, x>0$ is the kernel (part ignoring constants) of an exp(1) distribution

$$f_X(x) = e^{-x}, x>0$$

$$\int_0^{\infty} xe^{-x}dx = E[X] = 1$$

$$f_Y(y) = e^{-y}E[X] = e^{-y}$$

- The expectation of some variable can **never** be a function of that variable. If $E[Y]$ depends on the value of y, something has gone wrong.
- Similarly, if you want to find a marginal distribution, you integrate over the **other** variable. If you have a joint pdf $f_{X,Y}(x,y)$ and you want $f_X(x)$ it's, $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$. $f_X(x)$ should **never** depend on the value of Y.
- A fast way to lose points is not checking the bounds on your integrals and pdfs, either in calculations, or when reporting a final answer.
- If you are finding the pdf of a *function* of a variable (such as e^X), you must create a derived distribution – don't just plug in $f_X(e^X)$.
- If finding a normalization constant of a pdf (to make it integrate to 1), this will **always** be some number, not a function of the variables.