Problem solving strategies Part 2

Tara Templin

Review of Single-variable Derivation

Power Rule: Given: $f(x) = x^r$

First derivative w.r.t x: $f'(x) = r*x^{r-1}$

Derivatives are linear operators:

$$(f+g)'(x) = f'(x) + g'(x)$$

$$(cf)'(x) = cf'(x)$$

Product Rule:

$$(fg)' = f'g + fg'$$

Quotient Rule:

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

Review of Single-variable Derivation

Chain Rule: $f'(x) = h'(g(x)) \cdot g'(x)$

Examples of the Chain Rule in statistics:

$$\frac{\mathrm{d}}{\mathrm{dx}}[e^{f(x)}] = e^{f(x)} \cdot f'(x)$$

E.g. Normal, exponential distributions

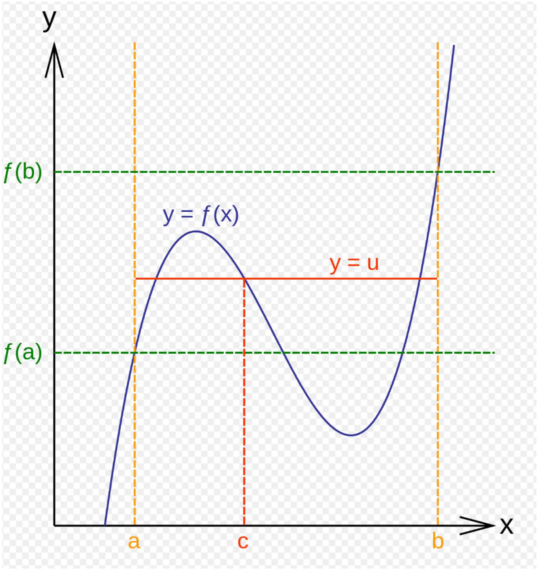
$$\frac{\mathrm{d}}{\mathrm{dx}}[\ln(f(x))] = \frac{1}{f(x)} \cdot f'(x)$$

E.g. log likelihood

Review of common calculus theorems

Intermediate value theorem: If f(x) is continuous on [a,b] and we define N as any number strictly between f(a) and f(b). Then there exists a number c in (a,b) such that f(c) = N.

Intuition: given f continuous on [1, 2] with the known values f(1) = 3 and f(2) = 5. Then the graph of y = f(x) must pass through the horizontal line y = 4 while x moves from 1 to 2.



https://en.wikipedia.org/wiki/Intermediate_value_theorem#/media/File:Intermediatevaluetheorem.svg

Review of common calculus theorems

Mean value theorem: If f(x) is continuous on [a,b] and differentiable on (a,b), then there is a number c in (a,b) such that f'(c) = (f(b)-f(a))/(b-a)

Extreme value theorem: If f'(x) is continuous on [a,b], then f(x) attains an absolute maximum value f(c) and an absolute minimum value f(d) at some numbers c and d in [a,b]

L'Hopital's Rule: $\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)}.$

(useful for 0/0 or infty/infty)

Review of Multivariable Differentiation

Joint PDF of X and Y: $f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$

$$\frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\left[F_{X,Y}(x, y)\right]\right] = \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}\left[F_{X,Y}(x, y)\right]\right] = f_{X,Y}(x, y)$$

Review of Single-variable Integrals

Integrals are linear operators:

$$\int (f(x) + g(x))dx = \int f(x)dx + \int g(x)dx$$

$$\int cf(x)dx = c \int f(x)dx$$

“Reverse” chain rule to solve integrals:

$$\int f(x)^n \cdot f'(x)\mathrm{dx} = f(x)^{n+1} \cdot \frac{1}{n+1}$$

Common integrals in statistics

$$\int e^{f(x)} \cdot f'(x)\mathrm{dx} = e^{f(x)}$$

$$\int \frac{f(x)}{f'(x)}\mathrm{dx} = \ln(|f(x)|)$$

Single-variable Integral Strategies

1) Simplify the integrand
   a) Do any obvious algebraic simplification of the integrand
   b) Use linearity if possible to break into easier integrals
   c) Try to manipulate the integrand to get f(x) and f'(x)
      i) Multiply by a/a where a is a constant, e^x/e^x, etc.
      ii) Multiply by the conjugate [e.g. if you have (x+5) multiply by (x-5)]
      iii) Complete the square
   d) Manipulate the integrand to make it look like a pdf, which integrates to 1
      i) Examples are Gamma and Beta PDFs - complex integrals which just equal 1!

$$\int f(x)^n \cdot f'(x)\mathrm{dx} = f(x)^{n+1} \cdot \frac{1}{n+1}$$

2) If all of above fails, look for u-substitution or solving with integration by parts.

Common tricks for solving integrals: PDFs

Uniform $\text{Unif}(a,b)$	$f(x) = \frac{1}{b-a}$ $x \in (a,b)$
Normal $\mathcal{N}(\mu,\sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$ $x \in (-\infty,\infty)$
Exponential $\text{Expo}(\lambda)$	$f(x) = \lambda e^{-\lambda x}$ $x \in (0,\infty)$
Gamma $\text{Gamma}(a,\lambda)$	$f(x) = \frac{1}{\Gamma(a)}(\lambda x)^a e^{-\lambda x}\frac{1}{x}$ $x \in (0,\infty)$
Beta $\text{Beta}(a,b)$	$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$ $x \in (0,1)$
Log-Normal $\mathcal{LN}(\mu,\sigma^2)$	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-(\log x-\mu)^2/(2\sigma^2)}$ $x \in (0,\infty)$
Chi-Square χ_n^2	$\frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}$ $x \in (0,\infty)$
Student-t t_n	$\frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$ $x \in (-\infty,\infty)$

All of the equations on the left integrate to 1 (by the definition of a PDF). Make sure the bounds of integral match the “support” (ie the “x ∈” stuff)

For any positive integer *n*, the gamma function:

$$\Gamma(n) = (n-1)!$$

is a constant. Treat it like a constant if it appears!

Use the definitions of “Gaussian integrals”

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}. \qquad \int_{-\infty}^{\infty} e^{-a(x+b)^2}\,dx = \sqrt{\frac{\pi}{a}}.$$

Source: http://www.wzchen.com/probability-cheatsheet/

Common tricks for solving integrals: odd/even

Check if you have an odd or an even function:

$$\int_{-t}^{t} o(x)\,dx = 0$$

$$\int_{-t}^{t} e(x)\,dx = 2\int_{0}^{t} e(x)\,dx$$

Check to see if the Integral is equal to the negative of itself:

$$\begin{aligned}\int_{-1}^{1} \frac{x^3-2x}{\sqrt{x^4+1}}\,dx &= \int_{-1}^{1} \frac{(-x)^3-2(-x)}{\sqrt{(-x)^4+1}}\,d(-x)\\ &= -\int_{-1}^{1} \frac{x^3-2x}{\sqrt{x^4+1}}\,dx\\ &= 0.\end{aligned}$$

https://brilliant.org/wiki/even-and-odd-functions/

Multiple Integrals

All but one variable is treated as constant when a partial derivative or integral is taken of a multivariate function.

If you have a function f(x,y) and you want the double integral over x and y, look for two things: (1) How to define your bounds and (2) Is it beneficial to swap the order of integration?

(1) “Setting bounds in terms of each other” = “Double Integrals over general regions”

Fubini's Theorem example

Compute $\int_0^1 \int_y^1 e^{-x^2} dx\, dy.$

We can't integrate because we don't have x. $\int e^{f(x)} \cdot f'(x)\mathrm{dx} = e^{f(x)}$

We can change the order of integration. $\int_0^1 \int_0^x e^{-x^2} dy\, dx = \int_0^1 x e^{-x^2} dx = \frac{1}{2}(1 - e^{-1}).$

Fubini's Theorem: If the integral is finite, you can interchange the order of integration

Notice we also had to adjust the bounds. We have to integrate over the same "area" ...Used to be across the x-axis, from x=y to x=1. Now it's up from the x-axis, y=0 to y=x and then across the y-axis, x=0 to x=1.

https://brilliant.org/wiki/multiple-integral/

Identifying functions of a random variable

1. X is the number of bikes you see in an hour.
2. 2X is the number of bike wheels you see in an hour.
3. X choose 2 = [x*(x-1)]/2 is the number of pairs of bikes such that you see both those bikes in that hour.
4. Integrals, Derivatives, and Summations are **linear operators/functions** on random variables
5. E[Y | X], the expected value of Y (number of graduate students on bikes) given the random variable X (number of bikes you see in an hour)

A function of a random variable is also a random variable. We can do all of the RV stuff (finding probability it falls in a range, expectations, variance, etc).

Source: http://www.wzchen.com/probability-cheatsheet/

Expected value of Y given the random variable X

E[Y | X] is **a function of the random variable X**. It is not a number except in certain cases such as if X and Y are independent.

In general, to find E[Y | X], find E[Y | X = x] and then plug in X for x.

For example: Say we are given X ~ N(0, 1) and Y = X^2.

E[Y | X = x] = x^2 since if we are given X = x then we know Y = x^2. E[Y | X] = X^2

E[X | Y = y] = 0 since if we know Y = y then we know X = ± sqrt(y), with equal probabilities (by symmetry of standard normal). E[X | Y] = 0

Source: http://www.wzchen.com/probability-cheatsheet/

Application: Surveys

Population mean E[X] and variance Var(X)

Population size, N

The sample size, n

The sample mean, max, median

The variance of the sampling distribution of mean

The estimated variance of the sample mean

Application: Surveys

Population mean E[X] and variance Var(X) = a number, not a RV

Population size, N = a number, not a RV

The sample size, n = usually a number, not a RV unless they give n ~ Distribution

The sample mean, max, median = RV

The variance of the sampling distribution of mean = a number σ^2/n. sigma is the pop var.

The estimated variance of the sample mean = RV

Conditional v Marginal Distributions

Conditional PDF of Y given X: The distribution of one random variable given certain restrictions for the other random variable.

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$

$$f_X(x|A) = \frac{P(A|X=x)f_X(x)}{P(A)}$$

Marginal PDF of X from Joint PDF: the distribution of one random variable without any sort of reference to the other random variable.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

http://www.wzchen.com/probability-cheatsheet/

Tip #1: If you have a series of steps, write out the overall plan before you get into the math

Conditional distribution of X given Y, $f_{X|Y}$

Conditional distribution of Y given X, $f_{Y|X}$

Joint distribution of X and Y, $f_{X,Y}$

Integrate over all values of y

Integrate over all values of x

Marginal distribution of X, f_X

Marginal distribution of Y, f_Y

Tip #2: Work in PDF space when asked for conditionals, marginals, joints. CDFs require extra steps

Conditional distribution of X given Y, $f_{X|Y}$

Conditional distribution of Y given X, $f_{Y|X}$

Divide

Joint distribution of X and Y, $f_{X,Y}$

Divide

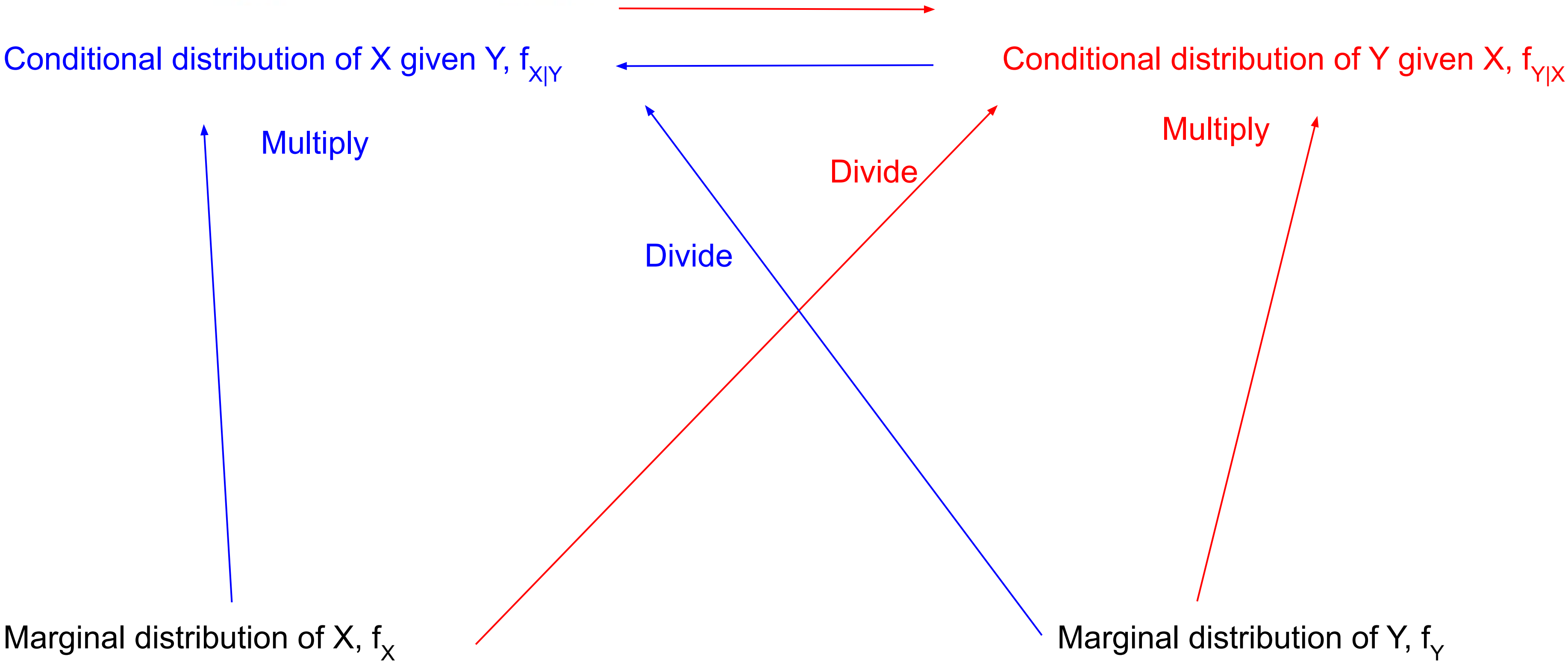
$$
\begin{aligned}
F_{X|Y}(x \mid y) &= \int_{-\infty}^{x} f_{X|Y}(t \mid y)\,\mathrm{d}t \\
&= \int_{-\infty}^{x} \frac{f_{X,Y}(t,y)}{f_Y(y)}\,\mathrm{d}t \\
&= \frac{1}{f_Y(y)} \int_{-\infty}^{x} f_{X,Y}(t,y)\,\mathrm{d}t
\end{aligned}
$$

Marginal distribution of X, f_X

Marginal distribution of Y, f_Y

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$



Tip #3: Look for independence. Random variables X and Y are independent **if and only if** any of the following conditions holds: (1) Joint CDF is the product of the marginal CDFs (2) Joint PMF/PDF is the product of the marginal PMFs/PDFs (3) Conditional distribution of Y given X is the marginal distribution of Y. Once you have one of these, you have all of them

Properties of Conditional Expectation

1. $E[Y \mid X] = E[Y]$ if X and Y are independent

2. $E[h(X)W \mid X] = h(X)*E[W \mid X]$ (taking out what's known)

   In particular, $E[h(X)|X] = h(X)$

3. $E[E[Y \mid X]] = E[Y]$

Law of Total Expectation: $E[Y] = E[Y \mid A] * Pr(A) + E[Y \mid A^c] * Pr(A^c)$

Law of Total Variance: $Var(Y) = E[Var(Y \mid X)] + Var(E[Y \mid X])$