14.102, Math for Economists
Fall 2005
Lecture Notes, 10/4/2005

These notes are primarily based on those written by Andrei Bremzen for 14.102 in 2002/3, and by Marek Pycia for the MIT Math Camp in 2003/4. I have made only minor changes to the order of presentation, and added a few short examples, mostly from Rudin. The usual disclaimer applies; questions and comments are welcome.

Nathan Barczi
nab@mit.edu

8 Calculus: Basic Integration in R

A function on a bounded interval $[a, b]$ is piece-wise continuous if it is continuous everywhere except on a finite number of points in I and that at every point where it is not continuous it admits finite left and right limits.

Definition 202 *For piecewise continuous functions on interval I, we define for any a and b in I $(a < b)$:*

$$\int_a^b f(x)dx = \lim_{n \to +\infty} \frac{b-a}{n} \sum_{k=1}^{n} f(a + \frac{k}{n}(b-a))$$

By definition: $\int_b^a f(x)dx = -\int_a^b f(x)dx$

The geometric interpretation of this integral (the Riemann integral) is the area under the curve. Look at the notation. Can show that such a limit exists. There are problems with this notion of the integral; too often it is not well defined.

It is possible to extend the definition of Riemann integral to a broader class of functions. One such useful generalization is Lebesgue integral based on measure theory. First note that measure theory generalizes the notions of length, area, and volume; a dose of it is helpful for studying probability theory for statistics and econometrics. Dividing $[a, b]$ into any measurable sets and using measures of set instead of lengths of intervals we can reproduce the above definition and basically construct the Lebesgue integral. The payoff from this more elaborate construction is that the limit "often" exist, and hence any "well-behaved" function is measurable and admits an integral (you may encounter some measurability problems in stochastic dynamic programming though).

Note, an important property of the Lebesgue integral is that the integral over sets of measure 0 has value 0 which is not necessarily the case for the Riemann construction! Otherwise Lebegue's has pretty much the same properties as

Riemann's (and it gives the same number as Riemann's whenever Riemann's exists).

A key example — probability density $f(x)$ and cumulative distribution $F(x)$.

Proposition 203 *Useful properties of the integral:*

- *if* $f \geq 0$ *and* $b \geq a$*, then* $\int_a^b f \geq 0$
- $\int_a^b (f+g) = \int_a^b f + \int_a^b g$
- $\int \lambda f(x)dx = \lambda \int f(x)dx$
- $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$
- *if* $f(x) \leq g(x)$ *for all* $x \in [a,b]$ *then* $\int_a^b f(x)dx \leq \int_a^b g(x)dx$
- $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx$
- $|\int_a^b f(x)dx| \leq \sup|f(x)| \times |b-a|$
- $|\int_a^b f(x)g(x)dx| \leq \sup|f(x)| \int_a^b |g(x)|dx$

Useful primitives:

- $\int_a^b (1/x)dx = \ln(b) - \ln(a)$
- $\int_a^b \exp(x)dx = \exp(b) - \exp(a)$
- $\int_a^b t^\alpha dt = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}$

8.1 Fundamental Theorem of Calculus

(a) Let f be integrable on an interval I, and let $a \in I$. Let $F(x) = \int_a^x f(t)dt$. If f is continuous at x then F is differentiable at x and

$$F'(x) \equiv \frac{d}{dx}\left[\int_a^x f(t)dt\right] = f(x).$$

(b) Suppose that F is differentiable on an interval I and that $F' = f$ is integrable. Then

$$\int_a^b f(x)dx = F(b) - F(a) \text{ for } a, b \in I.$$

Proof.

(a)

$$F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{\int_a^{x+h} f(t)dt - \int_a^x f(t)dt}{h} = \lim_{h\to 0} \frac{\int_x^{x+h} f(t)dt}{h}$$

Since, f is continuous at x so for any $\varepsilon > 0$ if h is small enough then we have

$$|f(x+h) - f(x)| < \varepsilon$$

and thus, if we define x^* as $x^* = x : |f(x^*) - f(x)| = \sup_{x' \in [x, x+h]} |f(x') - f(x)|$, we have

$$\left| \frac{F(x+h) - F(x)}{h} - \frac{hf(x)}{h} \right| = \left| \frac{\int_x^{x+h} f(t)dt}{h} - \frac{hf(x)}{h} \right| \leq \left| \frac{hf(x^*)}{h} - \frac{hf(x)}{h} \right| < \frac{h\varepsilon}{h} = \varepsilon$$

as $h \to 0$ we can take $\varepsilon \to 0$ and thus

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

(b) Let's consider only the easy case when f is continuous. Denote $G(x) = \int_a^x f(t)dt$ and notice that by (a) $G' = f$. Hence, F and G have same derivative, and $(F - G)' = 0$ and thus $F - G$ is constant, denote it c. Thus,

$$\begin{aligned} F(b) - F(a) &= (G(b) + c) - (G(a) + c) \\ &= G(b) - G(a) = \int_a^b f(t)dt. \end{aligned}$$

8.2 Change of variables and Integration by parts

The following two immediate consequences of the fundamental theorem of calculus are useful integration tools.

Theorem 204 *Change of variables: Let J_1 and J_2 be intervals (with more than one point):Let $f : J_1 \to J_2$ and $g : J_2 \to R$ continuous. Assume that f is differentiable and f' continuous. Then for any a, b in J_1*

$$\int_a^b g(f(x))f'(x)dx = \int_{f(a)}^{f(b)} g(u)du \tag{14}$$

Proof. Let $G' = g$. Then $(G \circ f)' = g(f(x))f'(x)$ by the chain rule. Both terms in (14) are equal to $G(f(b)) - G(f(a))$. ■

Example 205 $I = \int_0^1 2xe^{x^2} dx$, *say* $u = e^{x^2}$ *and* $du = 2xe^{x^2}$, $I = \int_1^e du = e - 1$.

Example 206 *Let us show the following, useful for moving from a normal to a standard normal distribution: If $f(x)$ is any pdf and μ and $\sigma > 0$ are any given constants, then the function $g(x|\mu, \sigma) = \frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$ is a pdf.*

To verify this, we must check that for all values of μ and σ, $g(x|\mu, \sigma)$ is nonnegative and integrates to 1. That it is nonnegative follows immediately from the fact that $f(x)$ is itself a pdf, and therefore is nonnegative for all values *of x (including $\frac{x-\mu}{\sigma}$). To check that g integrates to 1, we use the change of variables $y = \frac{x-\mu}{\sigma}, dy = \frac{1}{\sigma}dx$ to write*

$$\int_{-\infty}^{\infty} \frac{1}{\sigma} f(\frac{x-\mu}{\sigma})dx = \int_{-\infty}^{\infty} f(y)dy = 1$$

since $f(y)$ is a pdf.

Be familiar with the following two important special cases:

$$\int_a^b f(t+\alpha)dt = \int_{a+\alpha}^{b+\alpha} f(u)du$$

$$\int_a^b f(\alpha t)dt = \int_{\alpha a}^{\alpha b} \frac{f(u)}{\alpha}du$$

Proposition 207 ***Integration by parts:*** *Suppose F and G are differentiable on $[a,b]$. Suppose $F' = f$ and $G' = g$. are continuous. Then:*

$$\int_a^b f(t)G(t)dt = [F(b)G(b) - F(a)G(a)] - \int_a^b F(t)g(t)dt$$

You can obtain this formula quickly by noticing that $[FG]' = fG + Fg$.

Exercise 208 *Find $\int_a^x \log(t)dt$.*

8.3 Differentiation Under the Integral Sign

Often we encounter situations under which we wish to interchange the order of integration and differentiation.

Proposition 209 ***Leibniz rule:*** *If $f(t,x)$, $a(x)$, and $b(x)$ are differentiable with respect to x, then:*

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(t,x)dt = \int_{a(x)}^{b(x)} \frac{\partial f(t,x)}{\partial x}dt + b'(x)f(b(x),x) - a'(x)f(a(x),x)$$

Note that if $a(x)$ and $b(x)$ are constant, we have a special case:

$$\frac{d}{dx}\int_a^b f(t,x)dt = \int_a^b \frac{\partial f(t,x)}{\partial x}dt$$

Notice that this question really comes down to when it is justifiable to exchange the order of integration and a limit, since the derivative is a particular kind of limit. A full treatment of this question requires a bit of measure theory, which we won't go into here. However, a couple important results can be presented, all of which are variations on Lebesgue's Dominated Convergence Theorem (see Rudin; see also section 2.4 of Casella and Berger).

Theorem 210 *Suppose the function $h(x,y)$ is continuous at y_0 for each x, and there exists a function $g(x)$ satisfying*

1. $|h(x,y)| \leq g(x)$ *for all* x *and* y,
2. $\int_{-\infty}^{\infty} g(x)dx < \infty$

Then

$$\lim_{y \to y_0} \int_{-\infty}^{\infty} h(x,y)dx = \int_{-\infty}^{\infty} \lim_{y \to y_0} h(x,y)dx$$

The key condition is the existence of a dominating function $g(x)$, with a finite integral, which ensures that the integral of $h(x,y)$ cannot be too badly behaved. If we apply this to the case we are interested in, the derivative, we have

Theorem 211 *Suppose $f(t,x)$ is differentiable at $x = x_0$, that is,*

$$\lim_{h \to 0} \frac{f(t, x_0 + h) - f(t, x_0)}{h} = \left. \frac{\partial}{\partial x} f(t,x) \right|_{x=x_0}$$

exists for every t, and there exists a function $g(t, x_0)$, for all t and a constant $h_0 > 0$ such that

1. $\left| \frac{f(t,x_0+h)-f(t,x_0)}{h} \right| \leq g(t, x_0)$, *for all* t *and* $|h| \leq h_0$,
2. $\int_{-\infty}^{\infty} g(t, x_0)dx < \infty$.

Then

$$\left. \frac{d}{dx} \int_{-\infty}^{\infty} f(t,x)dx \right|_{x=x_0} = \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial x} \left. f(t,x) \right|_{x=x_0} \right] dx$$

The conditions essentially bound variability in the derivative of the function; they are similar to a smoothness condition called the *Lipschitz condition.* Most of the applications of these results which you'll see will come in statistics and econometrics, where many results in asymptotic theory examining the convergence behavior of a function as our data become infinite begin with a condition bounding the variance of the function in question.

Note that the theorem is stated for a particular value of x; often we have functions which are differentiable over some interval, and the theorem holds for x within this interval instead of a single value of x:

Theorem 212 *Suppose $f(t,x)$ is differentiable in x and there exists a function $g(t,x)$ and a constant h_0 such that*

1. $\left| \frac{\partial}{\partial x} f(t,x) \right|_{x=x'} \leq g(t,x)$ *for all* x' *such that* $|x' - x| \leq h_0$
2. $\int_{-\infty}^{\infty} g(t,x)dt < \infty$.

Then

$$\frac{d}{dx}\int_{-\infty}^{\infty} f(t,x)dt = \int_{-\infty}^{\infty} \frac{\partial}{\partial x} f(t,x)dt$$

Example 213 (moment generating functions, from Casella and Berger) *The moment generating function of a continuous random variable X is given by*

$$M_X(t) = \int_{-\infty}^{\infty} e^{tX} f_X(x)dx$$

Casella and Berger (Thm 2.3.7) tell us how to generate the moments of a distribution using the mgf:

$$EX^n = M_X^{(n)}(0) \equiv \frac{d^n}{dt^n} M_X(t)|_{t=0}$$

But the proof of this result assumes that we can interchange differentation and integration as follows:

$$\frac{d}{dt}M_X(t) = \frac{d}{dt}\int_{-\infty}^{\infty} e^{tX} f_X(x)dx = \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial t} e^{tX}\right) f_X(x)dx = E(Xe^{tX})$$

(we are focusing here on the first moment; plugging $t = 0$ into the last expression gives $\frac{d}{dt}M_X(t) = EX$, and the proof for higher moments continues in similar fashion, making similar assumptions about the interchangability of integration and differentiation).

We will show that this is true explicitly for a normal distribution with mean μ and variance 1 (to simplify the computations). We have

$$M_X(t) = Ee^{tX} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-(x-\mu)^2/2} dx$$

Applying the previous theorem requires finding a function $g(x,t)$ with finite integral such that

$$\left| \frac{\partial}{\partial t} e^{tx} e^{-(x-\mu)^2/2}|_{t=t'} \right| \leq g(x,t) \text{ for all } t' \text{ such that } |t'-t| \leq h_0 \qquad (*)$$

We have

$$\left| \frac{\partial}{\partial t} e^{tx} e^{-(x-\mu)^2/2} \right| = \left| xe^{tx} e^{-(x-\mu)^2/2} \right| \leq |x| \, e^{tx} e^{-(x-\mu)^2/2}$$

Now, define $g(x,t)$ separately for $x \geq 0$ and $x < 0$:

$$g(x,t) = \begin{cases} |x| \, e^{(t-h_0)x} e^{-(x-\mu)^2/2} & \text{if } x < 0 \\ |x| \, e^{(t+h_0)x} e^{-(x-\mu)^2/2} & \text{if } x \geq 0 \end{cases}$$

By construction, this function satisfies (*). But we need to check that its integral is finite.

For $x \geq 0$,

$$g(x,t) = xe^{-(x^2 - 2x(\mu+t+h_0)+\mu^2)/2}$$

We can complete the square in the exponent, writing

$$\begin{aligned} x^2 - 2x(\mu+t+h_0) + \mu^2 &= x^2 - 2x(\mu+t+h_0) + (\mu+t+h_0)^2 - (\mu+t+h_0)^2 + \mu^2 \\ &= (x-(\mu+t+h_0))^2 + \mu^2 - (\mu+t+h_0)^2 \end{aligned}$$

and now we have, for $x \geq 0$,

$$g(x,t) = xe^{-(x-(\mu+t+h_0))^2/2} e^{-(\mu^2-(\mu+t+h_0)^2)/2}$$

but the last exponential function doesn't depend on x, *and so we know that* $\int_0^\infty g(x,t)dx$ *is equal to a constant multiplied by a function which can be bounded by the mean of a normal distribution with mean* $\mu+t+h_0$. *Because we know that a normal distribution has finite mean (proven in Casella and Berger, Ch. 3), and employing a symmetric argument to cover the 'rest' of the integral (where* $x < 0$*), we have shown that* $g(x,t)$ *has a finite integral, justifying our exchange of differentiation and integration in using the mgf to calculate moments.*

8.4 Improper Integrals

Remark: improper integral: if $\lim_{A\to+\infty} \int_a^A f(t)dt$ exists, we note it $\int_a^{+\infty} f(t)dt$

Example 214 $\int_a^{+\infty} e^{-rt}dt = [-\frac{e^{-rt}}{r}]_a^{+\infty} = \frac{e^{-ra}}{r}$

Exercise 215 *compute* $\int_a^{+\infty} te^{-rt}dt$ *(use an integration by parts)*

Exercise 216 *compute* $\int_0^{+\infty} e^{-\sqrt{t}}dt$ *(use the change of variable* $u = \sqrt{t}$*)*