Math 10B with Professor Stankova
Worksheet, Discussion #17; Thursday, 3/21/2019
GSI name: Roy Zhao

Improper Integrals and Histograms

Concepts

1. An **improper integral** is an integral where one or two of the bounds of integration are $\pm\infty$. If both bounds are $\pm\infty$, we first need to split up the integral at some finite point (usually we pick 0 for convenience). Then to compute a one-sided improper integrals, we first write it as a limit as $n \to \infty$ then compute the limit. If any improper integral is ∞, we say the integral does not exist (even in the case of two sided integrals we get $\infty - \infty$).

   For a histogram, the area of the rectangle represents the probability of lying inside that region. So, if there is a 10% chance of X being between 10 and 15, the height of the rectangle would be $0.1/(15-10) = 0.02$.

Example

2. Calculate $\displaystyle\int_{-\infty}^{\infty} e^{-|x|}dx$.

Solution: We split it up at 0 to get

$$\int_{-\infty}^{\infty} e^{-|x|}dx = \int_{-\infty}^{0} e^{-|x|}dx + \int_{0}^{\infty} e^{-|x|}dx$$

$$= \lim_{t\to-\infty}\int_t^0 e^x dx + \lim_{t\to\infty}\int_0^t e^{-x}dx = \lim_{t\to-\infty} e^x|_t^0 + \lim_{t\to\infty}(-e^{-x})|0^t$$
$$= 1 - 0 + 0 - (-1) = 2.$$

3. Suppose exam scores are given by the following data. Construct a histogram of the data with intervals $[40, 50), [50, 60), \ldots, [90-100)$.

Score	40-49	50-59	60-69	70-79	80-89	90-99
Frequency	3	2	8	4	2	1

Solution: There are a total of 20 scores so we take the probabilities of getting in each range (frequency/20) then divide by the width of the rectangle (10). This gives heights of $0.015, 0.01, 0.04, 0.02, 0.01, 0.005$.

Problems

4. True **FALSE** When calculating $\int_{-\infty}^{\infty} f(x)dx$, the final result depends on the a we choose to split it up as $\int_{-\infty}^{a} f(x)dx + \int_{a}^{\infty} f(x)dx$.

Solution: The answer does not depend on a.

5. **TRUE** False Since the function $f(x) = x$ is odd, we know that $\int_{-n}^{n} f(x)dx = 0$ for all integers n.

Solution: Since it is odd, we can think of the area left of 0 and the area right of 0 as canceling.

6. True **FALSE** Since the function $f(x) = x$ is odd, we know that $\int_{-\infty}^{\infty} f(x)dx = 0$ for all integers n.

Solution: When we split up the integral, we get $\int_0^\infty xdx$ and $\int_{-\infty}^0 xdx$ which give ∞ and $-\infty$ respectively. Thus, the integral diverges.

7. Compute $\int_{-\infty}^{\infty} 2xe^{-x^2} dx$.

Solution: We split it up to get

$$\int_{-\infty}^{\infty} 2xe^{-x^2} = \lim_{t\to-\infty} \int_t^0 2xe^{-x^2} dx + \lim_{t\to\infty} \int_0^t 2xe^{-x^2} dx = \lim_{t\to-\infty} \int_{t^2}^0 e^{-u} du + \lim_{t\to\infty} \int_0^{t^2} e^{-u} du$$

$$= \lim_{t\to-\infty} -e^0 + e^{-t^2} + \lim_{t\to\infty} -e^{-t^2} + e^0 = -1 + 1 = 0.$$

8. Draw a new histogram using the previous data except with intervals $[40, 60), [60, 80), [80, 100)$.

Solution: We just add the frequencies of the two score ranges then we need to now divide the probability by 20. So the heights should be $\frac{5}{400}, \frac{12}{400}, \frac{3}{400}$.

Continuous Probability

Concepts

9. When we deal with random variables that can take a continuum of values, we have to use PDFs as opposed to PMFs. In this case, the value of the PMF does not give you a probability of picking that value, but instead gives you a relative likelihood. So if $f(x) = 2f(y)$, where f is the PDF, you expect to get a value near x around twice as likely as you are to get a value near y. Another difference from PMFs is now to calculate probabilities, we must take the integral along the interval we are asking about. So $P(a \leq X \leq b) = \int_a^b f(x)dx$. The most important property of a PDF is $\boxed{\int_{-\infty}^{\infty} f(x)dx = 1}$.

A CDF is a function $F(x)$ where $F(x) = P(X \leq x)$, it tells us that probability of getting a value less than or equal to x. It is just defined as $F(x) = \int_{-\infty}^{x} f(x)dx$. It satisfies three important properties:

- $F(x)$ is nondecreasing. So if $x \leq y$, then $F(x) \leq F(y)$.
- $\lim_{x \to -\infty} F(x) = 0$.
- $\lim_{x \to \infty} F(x) = 1$.

Example

10. Suppose that the probability density function P that an atom emits a gamma wave with the PDF $P(t) = Ce^{-10t}$ for $t \geq 0$ and $P(t) = 0$ for $t < 0$. Find P and calculate the CDF associated with P.

Solution: We solve the differential equation for P. We have that

$$\int \frac{dP}{P} = \int -10dt \implies \ln P = -10t + C \implies P = Ce^{-10t}.$$

Now in order for this to be a PDF, we need its integral to be 1. Thus, we have that

$$\int_{-\infty}^{\infty} P(t) = \int_{0}^{\infty} Ce^{-10t} = \frac{-Ce^{-10t}}{10}|_0^\infty = \frac{C}{10}.$$

Hence $C = 10$ and $P(t) = 10e^{-10t}$.

To find the CDF, we take the integral, for $t \geq 0$, we have that

$$F(t) = \int_{-\infty}^{t} 10e^{-10t} = \int_{0}^{t} 10e^{-10t} = -e^{-10t}|_0^t = 1 - e^{-10t}.$$

11. For the above PDF, find the probability that a gamma wave is emitted from -1 seconds to 1 second.

Solution: This is just the integral from -1 to 1. But, the probability is 0 from -1 to 0 so we just need to compute $P(0 \leq X \leq 1)$ which is

$$\int_0^1 P(t)dt = F(1) = 1 - e^{-10}.$$

Problems

12. True **FALSE** Since the CDF is an antiderivative of the PDF, there are multiple CDFs for a given PDF (and they differ by a $+C$).

Solution: We choose the CDF that gives us $F(-\infty) = 0$, which is like an initial condition and fixes the antiderivative.

13. True **FALSE** The area underneath a CDF must be equal to 1.

Solution: There is no such requirement on the CDF, and in fact, the area will actually diverge.

14. True **FALSE** A PDF must be continuous.

Solution: This is false. Look at the example above for a counterexample.

15. True **FALSE** Let $P(x) = Cx^3$ for $-1 \leq x \leq 2$ and 0 otherwise. Since $\int_{-1}^{3} P(x)dx = C(16 - 1/4)$, setting $C = (16 - 1/4)^{-1}$ makes P into a PDF.

Solution: P cannot possibly be a PDF because it is negative at some places.

16. Let $P(x) = Cx^2(10 - x)$ on $0 \leq x \leq 10$ and 0 otherwise. Find C such that P is a PDF and its corresponding CDF. Find the probability that we choose a number between 0 and 1.

Solution: In order for P to be a PDF, it needs to have integral 1. So

$$\int_{-\infty}^{\infty} P(x)dx = \int_{0}^{10} Cx^2(10 - x)dx = \frac{2500C}{3}.$$

Hence, we have that $C = \frac{3}{2500}$.

The CDF is the integral

$$\int_{-\infty}^{x} P(t)dt = \int_{0}^{x} \frac{3t^2(10 - t)}{2500} = \frac{(40 - 3x)x^3}{10000},$$

for $0 \leq x \leq 10$ and 0 for $x \leq 0$ and 1 for $x \geq 10$.

The probability that we choose a number between 0 and 1 is just the integral

$$\int_{0}^{1} P(x)dx = \int_{0}^{1} \frac{3x^2(10 - x)}{2500}dx = \frac{37}{10000} = 0.37\%.$$

17. Let $P(x) = C(x - 1)(x + 1)$ on $-1 \leq x \leq 1$ and 0 otherwise. Find C such that P is a PDF and its corresponding CDF. Find the probability that we choose a number between 0 and 1.

Solution: We need

$$\int_{-\infty}^{\infty} P(x)dx = \int_{-1}^{1} C(x - 1)(x + 1)dx = \frac{-4}{3}C = 1.$$

Hence $C = \frac{-3}{4}$. The CDF is

$$F(x) = \int_{-\infty}^{x} P(t)dt = \int_{-1}^{x} -3/4(t-1)(t+1)dt = \frac{-x^3+3x+2}{4}.$$

for $-1 \leq x \leq 1$ and 0 for $x \leq -1$ and 1 for $x \geq 1$.

The probability that we choose a number between 0 and 1 is

$$F(1) - F(0) = \int_{0}^{1} P(x)dx = \int_{0}^{1} -3/4(x-1)(x+1)dx = \frac{1}{2}.$$

18. Let $P(x)$ satisfy $\frac{dP}{dx} = 2x$ for $0 \leq x \leq 1$ and $P(x) = 0$ otherwise. Find P such that it is a PDF and its corresponding CDF. Find the probability that we choose a number between 0 and 1.

Solution: We have that $P(x) = x^2 + C$ for $0 \leq x \leq 2$ and for this to be a PDF, we require that

$$\int_{-\infty}^{\infty} P(x)dx = \int_{0}^{1} x^2 + Cdx = \frac{1}{3} + C = 1.$$

Hence $C = \frac{2}{3}$. The CDF is

$$F(x) = \int_{-\infty}^{x} P(t)dt = \frac{x(x^2+2)}{3}$$

for $0 \leq x \leq 1$ and 0 for $x \leq 0$ and 1 for $x \geq 1$.

The probability that we choose a number between 0 and 1 is 1 because our PDF is only between 0 and 1.

19. Let $F(x) = \frac{x-1}{x+1}$ for $x \geq 1$ and 0 for $x \leq 1$. Show that F is a CDF. Find the PDF associated with it and the probability that we choose a number between 1 and 2.

Solution: This is a CDF because it is continuous since $F(1) = 0$ and $\lim_{x\to\infty} F(x) = 1$ and F is non-decreasing. The PDF is

$$f(x) = \frac{d}{dx}F(x) = \frac{2}{(x+1)^2}$$

for $x \geq 1$ and 0 for $x \leq 1$. The probability that we choose a number between 1 and 2 is

$$\int_{1}^{2} f(x)dx = F(2) - F(1) = \frac{1}{3}.$$

20. Find numbers A, B such that $A\arctan(x) + B$ is a CDF and find the PDF associated with it. Find the probability that we choose a number between 0 and 1.

Solution: We know that $\arctan(x)$ is nondecreasing and all we need is for the range to be $(0, 1)$. The original range is $(-\pi/2, \pi/2)$ and so letting $A = 1/\pi$ changes our range to $(-1/2, 1/2)$, and shifting up by $B = 1/2$ gives a range of $(0, 1)$. Thus, we have that $A = \frac{1}{\pi}$ and $B = \frac{1}{2}$. The PDF is

$$f(x) = \frac{d}{dx}F(x) = \frac{1}{\pi + \pi x^2}.$$

The probability that we choose a number between 0 and 1 is

$$\int_0^1 f(x)dx = F(1) - F(0) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}.$$

21. Let $F(x) = \ln x$ for $1 \leq x \leq a$ and $F(x) = 0$ for $x \leq 1$ and $F(x) = 1$ for $x \geq a$. Find a such that F is a continuous CDF and find the PDF associated with it. Find the probability that we choose a number between 1 and 2.

Solution: In order for this to be a CDF, we require need this to be nondecreasing so a is such that $\ln(a) = 1$ or $a = e$. The PDF is the derivative so $f(x) = \frac{1}{x}$ between 1 and e and equal to 0 otherwise. The probability that we choose a number between 1 and 2 is $\ln(2) - \ln 1 = \ln 2$.