CSE 312 – Section 6

Spring 2026

Review of Main Concepts

- **Continuous Random Variable**: A continuous random variable X has an uncountably infinite number of values and its cumulative distribution function $F_X(x) : \mathbb{R} \to \mathbb{R}$ is continuous everywhere.

- **Cumulative Distribution Function (cdf)**: For any random variable (discrete or continuous) X, the cumulative distribution function is defined as $F_X(x) = \mathbb{P}(X \le x)$. Notice that (1) this function must be monotonically nondecreasing: if $x < y$ then $F_X(x) \le F_X(y)$, because $\mathbb{P}(X \le x) \le \mathbb{P}(X \le y)$; (2) since probabilities are between 0 and 1, that $0 \le F_X(x) \le 1$ for all x, with $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$; (3) since $\mathbb{P}(X = k) = 0$ for some constant k if X is a continuous random variable, $\mathbb{P}(X < k) = \mathbb{P}(X \le k)$.

- **Probability Density Function (pdf or density)**: Let X be a continuous random variable. Then the probability density function $f_X(x) : \mathbb{R} \to \mathbb{R}$ of X is defined as $f_X(x) = \frac{d}{dx} F_X(x)$. Taking the integral of both sides, it means that $F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(t)\, dt$. It follows that $\mathbb{P}(a \le X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$ and that $\int_{-\infty}^{\infty} f_X(x) dx = 1$. From the fact that $F_X(x)$ is monotonically nondecreasing it follows that $f_X(x) \ge 0$ for every real number x. Note that $f_X(a) \ne \mathbb{P}(X = a)$, since $\mathbb{P}(X = a) = F_X(a) - F_X(a) = 0$ for all a. However, the probability that X is close to a is proportional to $f_X(a)$: for small δ, $\mathbb{P}\left(a - \frac{\delta}{2} < X < a + \frac{\delta}{2}\right) \approx \delta f_X(a)$.

- **i.i.d. (independent and identically distributed)**: Random variables $X_1, \dots, X_n$ are i.i.d. (or iid) if they are independent and have the same probability mass function or probability density function.

- **Discrete to Continuous:** To summarize, when going from discrete to continuous, the main differences are (1) using an integral instead of a summation, and (2) using the density function $f_X(k)$ instead of the PMF $\mathbb{P}(X = k)$.

	Discrete	**Continuous**
PMF/PDF	$p_X(x) = \mathbb{P}(X = x)$	$f_X(x) \ne \mathbb{P}(X = x) = 0$
CDF	$F_X(x) = \sum_{t \le x} p_X(t)$	$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$
Normalization	$\sum_x p_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
Expectation	$\mathbb{E}[X] = \sum_x x p_X(x)$	$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$
LOTUS	$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

- **Standardizing**: Let X be any random variable (discrete or continuous, not necessarily normal), with $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$. If we let $Y = \frac{X-\mu}{\sigma}$, then $\mathbb{E}[Y] = 0$ and $\mathrm{Var}(Y) = 1$.

- **Law of Total Probability (Continuous): This may not have been covered in class yet, but will be at some point, and you will use it on the problem set.** A is an event, and X is a continuous random variable with density function $f_X(x)$.

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A \mid X = x) f_X(x) dx$$

- **Transforming Continuous Random Variables** (May not be covered in class.) Suppose that X is a discrete random variable that takes values in Ω_X and let $Y = g(X)$ for some function g. Let $\Omega_Y = \{g(x) | x \in \Omega_X\}$ Then the probability mass function of Y satisfies

$$p_Y(y) = \sum_{x \in \Omega_X | g(x) = y} p_X(x).$$

However, if X is a continuous random variable with density function f, and $Y = g(X)$ for some continuous function g, then we can **not** say that

$$f_Y(y) = \int_{x \in \Omega_X | g(x) = y} f_X(x) dx.$$

Rather, we must take the following steps:

  a) Compute $F_Y(y)$ from $F_X(x)$.
  b) Differentiate $F_Y(y)$ with respect to y to obtain $f_Y(y)$.

See Problems 17 and 18.

- **Zoo of *Continuous* Random Variables**

  a) **Uniform**: $X \sim \mathsf{Uniform}(a, b)$ iff X has the following probability density function:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

  $\mathbb{E}[X] = \frac{a+b}{2}$ and $\text{Var}(X) = \frac{(b-a)^2}{12}$. This represents each real number from $[a, b]$ to be equally likely.

  b) **Exponential**: $X \sim \mathsf{Exponential}(\lambda)$ iff X has the following probability density function:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

  $\mathbb{E}[X] = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$. $F_X(x) = 1 - e^{-\lambda x}$ for $x \geq 0$. The exponential random variable is the continuous analog of the geometric random variable: it represents the waiting time to the next event, where $\lambda > 0$ is the average number of events per unit time. Note that the exponential measures how much time passes until the next event (any real number, continuous), whereas the Poisson measures how many events occur in a unit of time (nonnegative integer, discrete). The exponential random variable X is memoryless:

$$\text{for any } s, t \geq 0,\ \mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$$

  The geometric random variable also has this property.

c) **Normal (Gaussian, "bell curve")**: $X \sim \mathcal{N}(\mu,\ \sigma^2)$ iff X has the following probability density function:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \quad x \in \mathbb{R}$$

$\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$. The "standard normal" random variable is typically denoted Z and has mean 0 and variance 1. The CDF has no closed form, but we denote the CDF of the standard normal as $\Phi(z) = F_Z(z) = \mathbb{P}(Z \leq z)$. Note from symmetry of the probability density function about $z = 0$ that: $\Phi(-z) = 1 - \Phi(z)$.

To find the values of $\Phi(\cdot)$, you can use this Z-table.

Closure of the Normal Distribution: Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then, $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. That is, linear transformations of normal random variables are still normal. Thus, for example, if $X \sim \mathcal{N}(\mu,\ \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$.

"Reproductive" Property of Normals: Let $X_1, \ldots, X_n$ be independent normal random variables with $\mathbb{E}[X_i] = \mu_i$ and $\text{Var}(X_i) = \sigma_i^2$. Let $a_1, \ldots, a_n \in \mathbb{R}$ and $b \in \mathbb{R}$. Then,

$$X = \sum_{i=1}^{n}(a_i X_i + b) \sim \mathcal{N}\left(\sum_{i=1}^{n}(a_i\mu_i + b), \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$$

There's nothing special about the parameters – the important result here is that the resulting random variable is still normally distributed.

- **Central Limit Theorem (CLT)**: Let $X_1, \ldots, X_n$ be iid random variables with $\mathbb{E}[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$. Let $X = \sum_{i=1}^{n} X_i$, which has $\mathbb{E}[X] = n\mu$ and $\text{Var}(X) = n\sigma^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, which has $\mathbb{E}[\overline{X}] = \mu$ and $\text{Var}(\overline{X}) = \frac{\sigma^2}{n}$. $\overline{X}$ is called the *sample mean*. Then, as $n \to \infty$, $\overline{X}$ approaches the normal distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Standardizing, this is equivalent to $Y = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ approaching $\mathcal{N}(0,1)$. Similarly, as $n \to \infty$, X approaches $\mathcal{N}(n\mu, n\sigma^2)$ and $Y' = \frac{X-n\mu}{\sigma\sqrt{n}}$ approaches $\mathcal{N}(0,1)$. It is no surprise that $\overline{X}$ has mean μ and variance σ^2/n – we have seen this before and it is easy to show. The importance of the CLT is that, for large n, regardless of what distribution X_i comes from, $\overline{X}$ is *approximately normally distributed with mean* μ *and variance* σ^2/n.

Announcements & Plan for Section

Announcements

- PSet 4 grades were released and can be viewed on Gradescope. Regrade requests will close on 5/9. We highly recommend taking a look at any feedback received, the common errors doc on Ed, and the solutions that were posted on Ed.
- Pset 5 was due yesterday.
- Pset 6 is released - will be due 2 weeks from now on 5/20.
- This week's focus: continuous distributions and midterm prep

Plan for Section

- Content Review (Problem 1)
- Go over practice midterm (linked below)

Suggested midterm problems to focus on: Task 1e, Task 2 (all parts), Task 3 c-f

Be sure to check out the remaining problems (especially 4, 10, and 14) before you do your homework.

Midterm Prep Resources

- Link to information about exam.
- Link to draft cheat sheet.
- Link to practice midterm and solutions to practice midterm

1 Content Review

a) What is $\mathbb{P}(X = 4)$ if X is a **continuous** random variable?

○ 1
○ 0
○ not enough information

b) The cumulative distribution function for a continuous random variable X is $F_X(k) =$

○ $\int_{-\infty}^{k} f_X(x)dx$
○ $\int_{-\infty}^{\infty} f_X(x)dx$
○ $\int_{k}^{\infty} f_X(x)dx$
○ $\frac{d}{dk} f_X(k)$

c) The probability density function for a continuous random variable X is $f_X(k) =$

○ $\int_{-\infty}^{k} f_X(x)dx$
○ $\frac{d}{dk} F_X(k)$

d) **True** or **False**. If X is a continuous random variable, $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)dx$

e) **True** or **False**. If X is a continuous random variable, $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

f) Which of the following follow an $\mathsf{Exponential}(\lambda)$ distribution?

○ Number of minutes to the first success with λ as average number of successes per minute
○ Number of successes in the first 1 minute with λ as average number of successes per minute
○ Time (real number) to the first success with λ as average number of successes per minute

g) True or False: For any random variable X, $\mathbb{P}(X = 5) = \mathbb{P}(X - 5 = 0)$.

h) True or False: For some continuous random variable X, $\mathbb{P}(X \leq 5) \neq \mathbb{P}(X < 5)$.

i) True or False: Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $a, b \in \mathbb{R}$. Then $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

2 Uniform2

Robbie decided he wanted to create a "new" type of distribution that will be famous, but he needs some help. He knows he wants it to be continuous and have uniform density, but he needs help working out some of the details. We'll denote a random variable X having the "Uniform-2" distribution as $X \sim \mathsf{Uniform2}(a, b, c, d)$, where $a < b < c < d$. We want the density to be non-zero in $[a, b]$ and $[c, d]$, and zero everywhere else. Anywhere the density is non-zero, it must be equal to the same constant.

a) Find the probability density function, $f_X(x)$. Be sure to specify the values (in terms of a, b, c, d) it takes on for every point in $(-\infty, \infty)$. (Hint: use a piecewise definition).

b) Find the cumulative distribution function, $F_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$. (Hint: use a piecewise definition).

3 Create the distribution

Suppose X is a continuous random variable that is uniform on $[0, 1)$ and uniform on $[1, 2]$, but

$$\mathbb{P}(1 \leq X \leq 2) = 2 \cdot \mathbb{P}(0 \leq X < 1).$$

Outside of $[0, 2]$ the density is 0. What is the PDF and CDF of X?

4 The Spotlight

A spotlight is mounted on a wall at a height h above the ground. The light rotates and is equally likely to point at any angle Θ between 0 and $\pi/4$ (where 0 corresponds to pointing straight down at the ground). Let X be the distance along the ground from the point directly beneath the light to the spot where the light hits the ground. (Note that $X = h\tan(\Theta)$). For X, find ...

a) the cumulative distribution function F_X.
Hint: First, determine the range of possible values for X given the bounds on Θ. Then, use the definition $F_X(x) = \mathbb{P}[X \leq x]$ and substitute for X. Recall that if $\tan(x) = y$, then $x = \arctan(y)$.

b) the probability density function f_X.
Hint: Recall the chain rule for derivatives, and that $\frac{\mathrm{d}}{\mathrm{d}u}\arctan(u) = \frac{1}{1+u^2}$.

c) the expected value $\mathbb{E}[X]$.
Hint: To evaluate the integral $\int \frac{x}{h^2+x^2}\,\mathrm{d}x$, try using the substitution $u = h^2 + x^2$.

d) the variance $\mathrm{Var}(X)$.
Hint: To find $\mathbb{E}[X^2]$, you will need to integrate a fraction like $\frac{x^2}{h^2+x^2}$. Try adding and subtracting h^2 in the numerator to split the fraction into two easier pieces: $x^2 = x^2 + h^2 - h^2$.

5 Max of uniforms

Let $U_1, U_2, \ldots, U_n$ be mutually independent Uniform random variables on $(0, 1)$. As in the discrete case, independence of these random variables implies that

$$\mathbb{P}(U_1 \leq x_1, \ldots, U_n \leq x_n) = \mathbb{P}(U_1 \leq x_1) \cdots \mathbb{P}(U_n \leq x_n)$$

for any numbers $x_1, \ldots, x_n$. Find the CDF and PDF for the random variable $Z = \max(U_1, \ldots, U_n)$.

6 New PDF?

Alex came up with a function that he thinks could represent a probability density function. He defined the potential PDF for X as $f(x) = \frac{1}{1+x^2}$ defined on $[0, \infty)$. Is this a valid PDF? If not, find a constant c such that the PDF $f_X(x) = \frac{c}{1+x^2}$ is valid. Then find $\mathbb{E}[X]$. (Hints: $\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$, $\tan \frac{\pi}{2} = \infty$, and $\tan 0 = 0$.)

7 Throwing a dart

Consider the closed unit circle of radius r, i.e., $S = \{(x, y) : x^2 + y^2 \leq r^2\}$. Suppose we throw a dart onto this circle and are guaranteed to hit it, but the dart is equally likely to land anywhere in S. Concretely this means that the probability that the dart lands in any particular area of size A (that is entirely inside the circle of radius R), is equal to $\frac{\text{A}}{\text{Area of whole circle}}$. The density outside the circle of radius r is 0.

Let X be the distance the dart lands from the center. What is the CDF and pdf of X? What is $\mathbb{E}[X]$ and $\text{Var}(X)$?

8 A square dartboard?

You throw a dart at an $s \times s$ square dartboard. The goal of this game is to get the dart to land as close to the lower left corner of the dartboard as possible. However, your aim is such that the dart is equally likely to land at any point on the dartboard. Let random variable X be the length of the side of the smallest *square* B in the lower left corner of the dartboard that contains the point where the dart lands. That is, the lower left corner of B must be the same point as the lower left corner of the dartboard, and the dart lands somewhere along the upper or right edge of B. For X, find the CDF, PDF, $\mathbb{E}[X]$, and $\text{Var}(X)$.

9 Will the battery last?

Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with expectation 10,000 miles. If the owner wants to take a 5000 mile road trip, what is the probability that she will be able to complete the trip without replacing the battery, given that the car has already been used for 2000 miles on the road trip?

10 Batteries and exponential distributions

Let X_1, X_2 be independent exponential random variables, where X_i has parameter λ_i, for $1 \leq i \leq 2$. Let $Y = \min(X_1, X_2)$.

a) Show that Y is an exponential random variable with parameter $\lambda = \lambda_1 + \lambda_2$. Hint: Start by computing $\mathbb{P}(Y > y)$. Two random variables with the same CDF have the same pdf. Why?

b) What is $\mathbb{P}(X_1 < X_2)$? Use the law of total probability. The law of total probability hasn't been covered in class yet, but will be soon at which point it would be good to revisit this problem!

c) You have a digital camera that requires two batteries to operate. You purchase n batteries, labelled $1, 2, \ldots, n$, each of which has a lifetime that is exponentially distributed with parameter λ, independently of all other batteries. Initially, you install batteries 1 and 2. Each time a battery fails, you replace it with the lowest-numbered unused battery. At the end of this process, you will be left with just one working battery. What is the expected total time until the end of the process? Justify your answer.

d) In the scenario of the previous part, what is the probability that battery i is the last remaining battery as a function of i? (You might want to use the memoryless property of the exponential distribution that has been discussed.)

11 Grading on a curve

In some classes (not CSE classes) an examination is regarded as being good (in the sense of determining a valid spread for those taking it) if the test scores of those taking it are well approximated by a normal density function. The instructor often uses the test scores to estimate the normal parameters μ and σ^2 and then assigns a letter grade of A to those whose test score is greater than $\mu + \sigma$, B to those whose score is between μ and $\mu + \sigma$, C to those whose score is between $\mu - \sigma$ and μ, D to those whose score is between $\mu - 2\sigma$ and $\mu - \sigma$ and F to those getting a score below $\mu - 2\sigma$. If the instructor does this and a student's grade on the test really is normally distributed with mean μ and variance σ^2, what is the probability that student will get each of the possible grades A,B,C,D and F?

12 Normal questions

a) Let X be a normal random variable with parameters $\mu = 10$ and $\sigma^2 = 36$. Compute $\mathbb{P}(4 < X < 16)$.

b) Let X be a normal random variable with mean 5. If $\mathbb{P}(X > 9) = 0.2$, approximately what is $\text{Var}(X)$?

c) Let X be a normal random variable with mean 12 and variance 4. Find the value of c such that

$$\mathbb{P}(X > c) = 0.10.$$

13 Do it in Reverse

a) Let X be a normal random variable with parameters $\mu = 8$ and $\sigma^2 = 9$. Find x such that $\mathbb{P}(X \leq x) = 0.6$.

b) Lots of statistics (like standardized test scores or heights) use *percentiles* to give context to where outcomes fall in a distribution. The nth percentile marks the outcome at which $n\%$ of the data points are less than the outcome. Let Y be a normal random variable with parameters $\mu = 15$ and $\sigma^2 = 4$. What value y marks the 85th percentile? What value b marks the 15th percentile?

14 Bad Computer

Each day, the probability your computer crashes is 10%, independent of every other day. Suppose we want to evaluate the computer's performance over the next 100 days.

a) Let X be the number of crash-free days in the next 100 days. What distribution does X have? Identify $\mathbb{E}[X]$ and $\mathrm{Var}(X)$ as well. Write an exact (possibly unsimplified) expression for $\mathbb{P}(X \geq 87)$.

b) Approximate the probability of at least 87 crash-free days out of the next 100 days using the Central Limit Theorem. Use continuity correction.

   Important: continuity correction says that if we are using the normal distribution to approximate

   $$\mathbb{P}\left(a \leq \sum_{i=1}^{n} X_i \leq b\right)$$

   where $a \leq b$ are integers and the X_i's are i.i.d. **discrete** random variables taking all integer values, then, as our approximation, we should use

   $$\mathbb{P}(a - 0.5 \leq Y \leq b + 0.5)$$

   where Y is the appropriate normal distribution that $\sum_{i=1}^{n} X_i$ converges to by the Central Limit Theorem. The intuition here is that, to avoid a mismatch between discrete distributions (whose range is a set of integers) and continuous distributions, we get a better approximation by imagining that a discrete random variable, say W, is a continuous distribution with density function

   $$f_W(x) := p_W(i) \qquad \text{when } i - 0.5 \leq x < i + 0.5 \text{ and } i \text{ integer}$$

   For more details see pages 209-210 in the Tsun book.

15 Another continuous r.v.

The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & \text{when } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathbb{E}[X] = \frac{3}{5}$, find a and b.

16 Point on a line

A point is chosen at random on a line segment of length L. Interpret this statement and find the probability that the ratio of the shorter to the longer segment is less than $\frac{1}{4}$.

17 Transforming continuous random variables

The next few questions are designed to help you understand the issue with transforming continuous random variables.

Specifically let's explore why we cannot simply adapt the discrete formula $p_Y(y) = \sum_{x|g(x)=y} p_X(x)$ into an integral $f_Y(y) = \int_{x|g(x)=y} f_X(x)\,\mathrm{d}x$ for continuous random variables.

a) Suppose $X \sim \text{Unif}(-1,1)$ and $Y = X^2$. We want to find the density of Y at $y = 0.25$. If we blindly apply the incorrect formula $f_Y(0.25) = \int_{x|x^2=0.25} f_X(x)\,\mathrm{d}x$, what is the mathematical result of this specific integral?

  (a) $f_X(-0.5) + f_X(0.5) = 1$
  (b) 0
  (c) 0.5
  (d) 0.25

b) Suppose $X \sim \text{Unif}(0,1)$ and we apply the transformation $Y = 3X$. We know that Y is uniformly distributed over $(0,3)$, so its true density should be $f_Y(y) = 1/3$ for $y \in (0,3)$.

If a student mistakenly assumes they can just "move" the density from X to Y by setting $f_Y(y) = f_X(x)$ where $3x = y$, what incorrect density would they get for Y?

  (a) $f_Y(y) = 1$ for $y \in (0,3)$
  (b) $f_Y(y) = 3$ for $y \in (0,3)$
  (c) $f_Y(y) = 1/3$ for $y \in (0,1)$
  (d) $f_Y(y) = 0$ for $y \in (0,3)$

c) Based on the previous examples, why does using the CDF method ($F_Y(y) = \mathbb{P}[Y \le y]$) succeed where manipulating the PDF directly fails?

  (a) The CDF converts the continuous variable into a discrete variable before taking the derivative.
  (b) The CDF works over intervals (areas) rather than individual points, properly accounting for how the transformation stretches or squishes the probability mass.
  (c) The CDF method only works for linear functions, bypassing the need for integration.
  (d) Probability density functions cannot be evaluated at specific points; they are only defined at infinity.

18 Non-Monotonic Transformations

When a transformation is not strictly increasing or decreasing, we have to be extra careful when setting up the inequalities for the CDF. Let $X \sim \text{Unif}(-2,2)$ and let $Y = X^2$.

(a) What is the cumulative distribution function (CDF) of Y? Be sure to clearly state the range of possible values for Y.
(b) Derive the probability density function (pdf) of Y by taking the derivative of your answer from part (a).

19 Transformations

Suppose $X \sim \mathsf{Uniform}(0,1)$ has the continuous uniform distribution on $(0,1)$. Let $Y = -\frac{1}{\lambda}\ln X$ for some $\lambda > 0$.

a) What is Ω_Y?

b) First write down $F_X(x)$ for $x \in (0,1)$. Then, find $F_Y(y)$ on Ω_Y.

c) Now find $f_Y(y)$ on Ω_Y (by differentiating $F_Y(y)$ with respect to y). What distribution does Y have?