CSE 312 Section 5

Continuous Random Variables

Slides made by Harrison Bay

Administrivia

Announcements & Reminders

- HW3
  - Grades released on Gradescope – check your submission to read comments
  - Regrade requests open ~24 hours after grades are released and close after a week

- HW4
  - Due Tuesday 7/24 @ 11pm

Review & Questions

Review of Main Concepts

- **Cumulative Distribution Function (CDF)**: For any random variable X, the CDF is defined as $F_X(x) = \mathbb{P}(X \le x)$.
  - Notice the CDF is monotonic (non-decreasing), i.e. if $x < y$ then $F_X(x) \le F_X(y)$
  - The CDF is bounded between 0 and 1.
- A **Continuous RV** is one for which the CDF is continuous everywhere.
  - Support is an uncountably infinite set.
- **Probability Density Function (PDF)** is defined as $f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x)$
  - Implies that $F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t$
  - $\mathbb{P}(a \le X \le b) = \int_{a}^{b} f_X(t)\mathrm{d}t$
  - $f_X(x) \ge 0$

Review of Main Concepts

- **I.I.D**: Random variables are "independent and identically distributed" if they are independent and have the same PDF/PMF
- For continuous random variables:
  - $\mathbb{P}(X = x) = 0 \neq f_X(x)$
  - $F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t$
  - $\int_{-\infty}^{\infty} f_X(t)\mathrm{d}t = 1$
  - $\mathbb{E}[X] = \int_{-\infty}^{\infty} t \cdot f_X(t)\mathrm{d}t$
  - $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(t) \cdot f_X(t)\mathrm{d}t$

Review of Main Concepts – Distributions

Discrete Distributions

Distribution	Parameters	Possible Description	Range Ω_X	$\mathbb{E}[\mathbf{X}]$	$\mathrm{Var}(\mathbf{X})$	PMF ($p_X(k)$ for $k \in \Omega_X$)
Uniform (disc)	$X \sim \mathrm{Unif}(a,b)$ for $a,b \in \mathbb{Z}$ and $a \le b$	Equally likely to be any *integer* in $[a,b]$	$\{a,\dots,b\}$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+2)}{12}$	$\frac{1}{b-a+1}$
Bernoulli	$X \sim \mathrm{Ber}(p)$ for $p \in [0,1]$	Takes value 1 with prob p and 0 with prob $1-p$	$\{0,1\}$	p	$p(1-p)$	$p^k(1-p)^{1-k}$
Binomial	$X \sim \mathrm{Bin}(n,p)$ for $n \in \mathbb{N}$, and $p \in [0,1]$	Sum of n iid $Ber(p)$ rvs. # of heads in n independent coin flips with $P(head) = p$.	$\{0,1,\dots,n\}$	np	$np(1-p)$	$\binom{n}{k}p^k(1-p)^{n-k}$
Poisson	$X \sim \mathrm{Poi}(\lambda)$ for $\lambda > 0$	# of events that occur in **one** unit of time independently with rate λ per unit time	$\{0,1,\dots\}$	λ	λ	$e^{-\lambda}\frac{\lambda^k}{k!}$
Geometric	$X \sim \mathrm{Geo}(p)$ for $p \in [0,1]$	# of independent Bernoulli trials with parameter p *up to and **including** first success*	$\{1,2,\dots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$(1-p)^{k-1}p$
Hypergeometric	$X \sim \mathrm{HypGeo}(N,K,n)$ for $n, K \le N$ and $n,K,N \in \mathbb{N}$	# of successes in n draws (w/o replacement) from N items that contain K successes in total	$\{\max(0, n+K-N), \dots, \min(n,K)\}$	$n\frac{K}{N}$	$n\frac{K(N-K)(N-n)}{N^2(N-1)}$	$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$
Negative Binomial	$X \sim \mathrm{NegBin}(r,p)$ for $r \in \mathbb{N}, p \in [0,1]$	Sum of r iid $Geo(p)$ rvs. # of independent flips until r^{th} head with $P(head) = p$	$\{r, r+1, \dots\}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\binom{k-1}{r-1}p^r(1-p)^{k-r}$
Multinomial	$\mathbf{X} \sim \mathrm{Mult}_r(n,\mathbf{p})$ for $r,n \in \mathbb{N}$ and $\mathbf{p} = (p_1, p_2, \dots, p_r)$, $\sum_{i=1}^r p_i = 1$	Generalization of the Binomial distribution, n trials with r categories each with probability p_i	$k_i \in \{0,\dots,n\}$, $i \in \{1,\dots,r\}$ and $\Sigma k_i = n$	$\mathbb{E}[\mathbf{X}] = n\mathbf{p} = \begin{bmatrix} np_1 \\ \vdots \\ np_r \end{bmatrix}$	$Var(X_i) = np_i(1-p_i)$ $Cov(X_i, X_j) = -np_ip_j, i \ne j$	$\binom{n}{k_1,\dots,k_r}\prod_{i=1}^r p_i^{k_i}$
Multivariate Hypergeometric	$\mathbf{X} \sim \mathrm{MVHG}_r(N,\mathbf{K},n)$ for $r,n \in \mathbb{N}$, $\mathbf{K} \in \mathbb{N}^r$ and $N = \sum_{i=1}^r K_i$	Generalization of the Hypergeometric distribution, n draws from r categories each with K_i successes (w/out replacement)	$k_i \in \{0,\dots,K_i\}$, $i \in \{1,\dots,r\}$ and $\Sigma k_i = n$	$\mathbb{E}[\mathbf{X}] = n\frac{\mathbf{K}}{N} = \begin{bmatrix} n\frac{K_1}{N} \\ \vdots \\ n\frac{K_r}{N} \end{bmatrix}$	$Var(X_i) = n\frac{K_i}{N} \cdot \frac{N-K_i}{N} \cdot \frac{N-n}{N-1}$ $Cov(X_i, X_j) = -n\frac{K_i}{N}\frac{K_j}{N} \cdot \frac{N-n}{N-1}, i \ne j$	$\frac{\prod_{i=1}^r \binom{K_i}{k_i}}{\binom{N}{n}}$

Continuous Distributions

Distribution	Parameters	Possible Description	Range Ω_X	$\mathbb{E}[\mathbf{X}]$	$\mathrm{Var}(\mathbf{X})$	PDF $f_X(x)$ for $x \in \Omega_X$	CDF ($F_X(x) = \mathbb{P}(X \le x)$)
Uniform	$X \sim \mathrm{Unif}(a,b)$ for $a < b$	Equally likely to be any real number in $[a,b]$	$[a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{1}{b-a}$	$\begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x < b \\ 1 & \text{if } x \ge b \end{cases}$

Source: https://courses.cs.washington.edu/courses/cse312/23su/resources/distributions.pdf

Problem 1 – True or False

1 – True or False

Identify the following statements as true or false (true means always true). Justify your answer.

a) For any random variable X, we have $\mathbb{E}[X^2] \geq \mathbb{E}[\mathrm{X}]^2$.
b) Let X, Y be r.v.s. Then, X and Y are independent if and only if $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.
c) Let $X \sim \mathrm{Bin}(n, p)$ and $Y \sim \mathrm{Bin}(m, p)$ be independent. Then, $X + Y \sim \mathrm{Bin}(n + m, p)$.
d) Let $X_1, \dots, X_{n+1}$ be independent $\mathrm{Ber}(p)$ r.v.s. Then, $\mathbb{E}[\sum_{i=1}^{n} X_i X_{i+1}] = np^2$.
e) Let $X_1, \dots, X_{n+1}$ be independent $\mathrm{Ber}(p)$ r.v.s. Then, $Y = \sum_{i=1}^{n} X_i X_{i+1} \sim \mathrm{Bin}(n, p^2)$.
f) If $X \sim \mathrm{Ber}(p)$, then $nX \sim \mathrm{Bin}(n, p)$.
g) If $X \sim \mathrm{Bin}(n, p)$, then $\frac{X}{n} \sim \mathrm{Ber}(p)$.
h) For any random variables X, Y, we have $Var(X - Y) = Var(X) - Var(Y)$.

1 – True or False

Identify the following statements as true or false (true means always true). Justify your answer.

a) For any random variable X, we have $\mathbb{E}[X^2] \geq \mathbb{E}[\mathrm{X}]^2$. True

b) Let X, Y be r.v.s. Then, X and Y are independent if and only if $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. False

c) Let $X \sim \mathrm{Bin}(n, p)$ and $Y \sim \mathrm{Bin}(m, p)$ be independent. Then, $X + Y \sim \mathrm{Bin}(n + m, p)$. True

d) Let $X_1, \dots, X_{n+1}$ be independent $\mathrm{Ber}(p)$ r.v.s. Then, $\mathbb{E}[\sum_{i=1}^{n} X_i X_{i+1}] = np^2$. True

e) Let $X_1, \dots, X_{n+1}$ be independent $\mathrm{Ber}(p)$ r.v.s. Then, $Y = \sum_{i=1}^{n} X_i X_{i+1} \sim \mathrm{Bin}(n, p^2)$. False

f) If $X \sim \mathrm{Ber}(p)$, then $nX \sim \mathrm{Bin}(n, p)$. False

g) If $X \sim \mathrm{Bin}(n, p)$, then $\frac{X}{n} \sim \mathrm{Ber}(p)$. False

h) For any random variables X, Y, we have $Var(X - Y) = Var(X) - Var(Y)$. False

Problem 2 – Fun with Poissons

2 – Fun with Poissons

See full proof on solution pdf.

Problem 7 – Throwing a dart

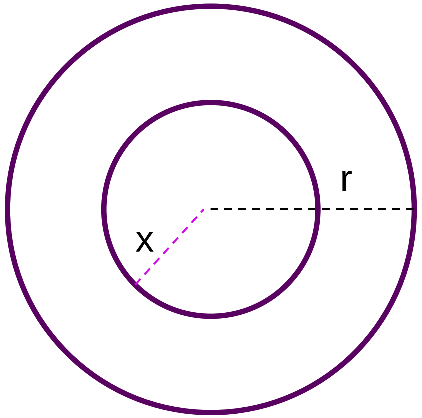
Problem 5

- 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 X. Concretely, this means that the probability that the dart lands in any particular area of size A is equal to $\frac{A}{Area\ of\ the\ whole\ circle}$. The density outside the unit circle is 0.
- Let X be the distance the dart lands from the center. What is the CDF and PDF of X? What is $E[X]$ and $Var(X)$?

Problem 5 Solution

- $F_X(x)$ is the probability that the dart lands inside the circle of radius x.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2/r^2 & 0 < x \leq r \\ 1 & x > r \end{cases}$$



Problem 5 Solution

- $F_X(x)$ is the probability that the dart lands inside the circle of radius x.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2/r^2 & 0 < x \le r \\ 1 & x > r \end{cases}$$

- To find the PDF we take the derivative of the CDF:

$$f_X(x) \begin{cases} 2x/r^2 & 0 < x \le r \\ 0 & otherwise \end{cases}$$

Problem 5 Solution

- $F_X(x)$ is the probability that the dart lands inside the circle of radius x.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2/r^2 & 0 < x \le r \\ 1 & x > r \end{cases}$$

- To find the PDF we take the derivative of the CDF:

$$f_X(x) \begin{cases} 2x/r^2 & 0 < x \le r \\ 0 & otherwise \end{cases}$$

- Use the definition of expectation:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^r \frac{2x^2}{r^2} dx = \frac{2}{3x^2}\left(x^3 \Big|_0^r\right) = \frac{2}{3} r$$

Problem 5 Solution

- $F_X(x)$ is the probability that the dart lands inside the circle of radius x.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2/r^2 & 0 < x \leq r \\ 1 & x > r \end{cases}$$

- To find the PDF we take the derivative of the CDF:

$$f_X(x) \begin{cases} 2x/r^2 & 0 < x \leq r \\ 0 & otherwise \end{cases}$$

- Use the definition of expectation:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^r \frac{2x^2}{r^2} dx = \frac{2}{3r^2}\left(x^3 \Big|_0^r\right) = \frac{2}{3} r$$

- Use $Var(X) = E[X^2] - E[X]^2$:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^r \frac{2x^3}{r^2} dx = \frac{2}{4r^2}\left(x^4 \Big|_0^r\right) = \frac{1}{2} r^2$$

$$\Rightarrow Var(X) = \frac{1}{2} r^2 - \left(\frac{2}{3} r\right)^2 = \frac{1}{18} r^2$$

Problem 3 – Uniform2

Problem 3

- 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 $X \sim Uniform2(a, b, c, d)$ 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 PDF $f_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$.
  - (b) Find the CDF, $F_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$.

Problem 3 Solution

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

Problem 3 Solution

- $f_X(x) = \begin{cases} \frac{1}{(b-a)+(d-c)}, & x \in [a,b] \cup [c,d] \\ 0, & otherwise \end{cases}$
- $F_X(x) = \begin{cases} 0, & x \in (-\infty, a) \\ \frac{(x-a)}{(b-a)+(d-c)}, & x \in [a,b) \\ \frac{(b-a)}{(b-a)+(d-c)}, & x \in [b,c) \\ \frac{(b-a)+(x-c)}{(b-a)+(d-c)}, & x \in [c,d) \\ 1, & x \in [d,\infty) \end{cases}$