Lecture 1. Math Review (1)

Sim, Min Kyu, Ph.D.
mksim@seoultech.ac.kr

Math Review 1 - Random Variables - pmf/pdf

- Class of a random variable (r.v.)
  - Discrete r.v.: Toss a coin, Throw a dice
  - Continuous r.v.: Uniform r.v., Normal r.v.

Definition - pmf (probability mass function)

- A pmf for a discrete r.v. X, $p(x)$, is a **function that gives the probability that a discrete random variable is exactly equal to some value**, i.e. $\mathbb{P}(X = x) = p(x)$.
- For example, suppose you throw a dice and X be a r.v. of the outcome. What is the pmf of X?

Definition - pdf (probability density function)

- A pdf for a continuous r.v. X, $f(x)$, is a **function that gives the relative likelihood for this continuous random variable to take on a given value**, i.e. $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$

Math Review 1 - Random Variables - pmf/pdf

- Properties of pmf/pdf[1]
  - The functions (pmf and pdf) are nonnegative everywhere.
    - i.e. $p(x) \geq 0$, $f(x) \geq 0$
  - Its summation/integral over the entire area is equal to one.
    - i.e. $\sum p(x) = 1$, $\int f(x)dx = 1$.
  - One can derive cdf (cumulative distribution function) from pmf and pdf.
    - (See the definition of cdf.)

1: A mathematical property is a something that always happen.

Math Review 2 - Random Variables - expectation

- For a discrete random variable X with pmf $p(x)$
  - $\mathbb{E}X = \sum xp(x)$
  - $\mathbb{E}[g(X)] = \sum g(x)p(x)$
    - ex) $\mathbb{E}X^2 = \sum x^2 p(x)$
- For a continuous random variable X with pdf $f(x)$
  - $\mathbb{E}X = \int_{-\infty}^{\infty} xf(x)dx$
  - $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
    - ex) $\mathbb{E}X^2 = \int_{-\infty}^{\infty} x^2 f(x)dx$

Math Review 3 - Random Variables - cdf

Definition - cdf (cumulative distribution function)

- For a random variable X, the cdf(cumulative distribution function) $F(x)$ **is a function of probability that the random variable X is found at a value less than or equal to x.**
  - $F(x) := \mathbb{P}(X \leq x)$
  - If discrete, $F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(-\infty \leq X \leq x) = \sum_{y=-\infty}^{y=x} \mathbb{P}(X = y) = \sum_{y=-\infty}^{y=x} p(y)$
  - If continuous, $F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(-\infty \leq X \leq x) = \int_{-\infty}^{x} f(y)dy$

blank

Math Review 4 - Uniform distribution

Definition - Uniform distribution

- A continuous random variable X is said to follow **Uniform distribution** with parameter a and b, and write $X \sim U(a, b)$ if

$$\text{pdf } f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$$\text{cdf } F(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

- Check: pdf $\rightarrow$ cdf

- For $X \sim U(a, b)$, show pdf $\rightarrow$ cdf

Math Review 5 - Exponential distribution

Definition - Exponential distribution

- A nonnegative continuous random variable X is said to follow an **exponential distribution** with parameter λ and write $X \sim exp(\lambda)$, if

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

$$\text{cdf } F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- TFAE (The followings are all equivalent)
  - $X \sim exp(\lambda)$
  - pdf $f(x) = \lambda e^{-\lambda x}$, if $x \geq 0$; $f(x) = 0$, otherwise
  - cdf $F(x) = 1 - e^{-\lambda x}$, if $x \geq 0$; $F(x) = 0$, otherwise
- pf) pdf $\rightarrow$ cdf

Math Review 6 - Random Variables - summary statistics

- Expectation: $\mathbb{E}X = \int_{-\infty}^{\infty} xf(x)dx$ or $\sum xp(x)$.
- Variance: $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}(X - \mathbb{E}X)^2$
- Standard Deviation: $sd(X) = \sqrt{Var(X)}$

- cv (coefficient of variation) $c_X = \frac{sd(X)}{\mathbb{E}X}$
- scv (squared-cv) $c_X^2 = \frac{(sd(X))^2}{(\mathbb{E}X)^2} = \frac{Var(X)}{(EX)^2}$

- Examples
  - ex) For an Exponential dist. $X \sim exp(\lambda)$, $c_X =$
  - ex) For a Normal dist. $X \sim N(\mu, \sigma^2)$, $c_X =$
  - ex) For deterministic X, $c_X =$

Math Review 7 - Exponential distribution - properties

- For $X \sim exp(\lambda)$, $\mathbb{E}X = 1/\lambda$, $Var(X) = 1/\lambda^2$, $sd(X) = 1/\lambda$
- Theorems
  - Memoryless property
  - Suppose $X_1 \sim exp(\lambda_1), X_2 \sim exp(\lambda_2)$, and they are independent, then $\mathbb{P}(X_1 < X_2) = \frac{\lambda_1}{\lambda_1+\lambda_2}$
  - If $X_1 \sim exp(\lambda_1), X_2 \sim exp(\lambda_2)$, and they are independent, then $min(X_1, X_2) \sim exp(\lambda_1 + \lambda_2)$

Math Review 8 - Exponential distribution - moments

- $\mathbb{E}X = 1/\lambda, Var(X) = 1/\lambda^2$

Math Review 9 - Memoryless property

Definition - Memoryless property

- A r.v. X is *memoryless*, if $\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s)$, for $s, t \geq 0$.

Theorem

- Exponential random variable is memoryless.

blank

Math Review 10 - Exponential distribution - property

Theorem

- Suppose $X_1 \sim exp(\lambda_1), X_2 \sim exp(\lambda_2)$, and they are independent, then $\mathbb{P}(X_1 < X_2) = \frac{\lambda_1}{\lambda_1+\lambda_2}$

Example

- Smith and Jones came to post office together and they are served by two clerks, A and B, respectively. Server A has service time following $exp(1/3)$ and server B has service time following $exp(1/5)$.
    1. What is the chance that Smith will be done first?
    2. Suppose that Smith came to post office earlier than Jones by 2 minutes, but Smith was still being served at the moment that Jones started to being served. Would this assumption change your previous answer?

Math Review 11 - Poisson distribution

Definition - Poisson distribution

- A discrete nonnegative random variable X is said to follow a **Poisson distribution** with parameter λ, and write $X \sim Poi(\lambda)$, if pmf

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \text{ for } k = 0, 1, 2, \ldots$$

- What is cdf of $Poi(\lambda)$?
- $\mathbb{E}X = Var(X) = \lambda$.

- We have
$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \text{ for } k = 0, 1, 2, ...$$
- $\mathbb{E}X = \lambda$
- $Var(X) = \lambda$

Math Review 12 - Probabilities

Definition - Probability

- Probability is a function from **"an event"** to real number between 0 and 1.

- $\mathbb{P} : S \to [0,1] \cap \mathbb{R}$
- $\mathbb{P}(S) = 1$
- For any event E, $0 \leq \mathbb{P}(E) \leq 1$
- For $E = \emptyset$, $\mathbb{P}(E) = 0$
- If $E_1 \cap E_2 = \emptyset$, then $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$.
- $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2)$.

Math Review 13 - Conditional Probabilities

Definition - Conditional Probability

- $\mathbb{P}(E|F)$ is the **probability that event E occurs given that F has occurred**.
- $\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$

Example. (from Ross)

- Suppose cards numbered one through ten are placed in a hat, mixed up, and then one of the cards is drawn. If we are told that the number on the drawn card is at least five, then what is the conditional probability that it is ten?

Math Review 14 - Probabilities - properties

- $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$, if $E_1 \cap E_2 = \emptyset$.
- $\mathbb{P}(E) = \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c) \because (E \cap F) \cap (E \cap F^c) = \emptyset$
- If $F_i \cap F_j = \emptyset$ for all $i \neq j$ and $\mathbb{P}(F_1 \cup \cdots \cup F_n) = 1$, then $\mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E \cap F_i)$

Math Review 15 - Bayes' rule

Definition - Bayes' rule

- Definition $F_1, \ldots, F_n$ are a mathematical partition if
  - $F_i \cap F_j = \emptyset$, or, $\mathbb{P}(F_i \cap F_j) = 0$, for any $i \neq j$
  - $\mathbb{P}(F_1 \cup \cdots \cup F_n) = 1$
- In other words, exactly one of the events $F_1, \ldots, F_n$ will occur.
- Then, the following holds.

$$\begin{aligned} \mathbb{P}(E) &= \mathbb{P}(E \cap F_1) + \cdots + \mathbb{P}(E \cap F_n) \\ &= \sum_{i=1}^{n} \mathbb{P}(E \cap F_i) \\ &= \sum_{i=1}^{n} \mathbb{P}(E|F_i)\mathbb{P}(F_i) \end{aligned}$$

Math Review 16 - Properties of max/min

1. $-max(x, y) = min(-x, -y)$
2. $max(x, y) = -(-max(x, y)) = -min(-x, -y)$
3. $max(x, 0) = -min(-x, 0)$
4. $max(x, y) + z = max(x + z, y + z)$
5. If x^* maximizes(minimizes) $f(x)$, then x^* minimizes(maximizes) $-f(x)$
6. If x^* maximizes(minimizes) $f(x)$, then x^* maximizes(minimizes) $f(x) + g(y) + C$.

```
"Man can learn nothing unless he proceeds from the known to the unknown.
- Claude Bernard"
```