# Review

Hyonho Chun

Department of Mathematical Sciences
KAIST

March 11, 2020

# Probability

- $S$: sample space
- Sigma algebra $\mathscr{B}$: a collection of subsets of $S$ that satisfies
  1. $\emptyset \in \mathscr{B}$
  2. If $A \in \mathscr{B}$, then $A^c \in \mathscr{B}$
  3. If $A_1, A_2, \ldots \in \mathscr{B}$, then $\cup_{i=1}^{\infty} A_i \in \mathscr{B}$.

  Show that $\cap_{i=1}^{\infty} A_i \in \mathscr{B}$.
- $P$ is a probability function if $P$ satisfies
  1. $P(A) \geq 0$ for all $A \in \mathscr{B}$
  2. $P(S) = 1$
  3. If $A_1, A_2, \ldots \in \mathscr{B}$ are pairwise disjoint, then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.
- Theorem

  If $P$ is a probability
  1. $P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$ for any partition $C_1, C_2, \ldots$
  2. $P(\cup_{i=1}^{\infty} A_i) \leq \sum P(A_i)$.

Proof)

- $S = \cup_{i=1}^{\infty} C_i$
- $P(A) = P(A \cap S) = P(A \cap \cup_{i=1}^{\infty} C_i) = P(\cup_{i=1}^{\infty}(A \cap C_i)) = \sum_{i=1}^{\infty} P(A \cap C_i)$
- Consider $B_i = A_i \backslash (A_1 \cup A_2 \cup \ldots \cup A_{i-1})$
- $P(\cup_{i=1}^{\infty} A_i) = P(\cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i) \leq \sum_{i=1}^{\infty} P(A_i)$

# Probability

- Conditional Probability $P(A|B) = P(A \cap B)/P(B)$.
- Example 1.3.4 (Three prisoners)
  Three prisoners, A, B, and C, are on death row. The governor decides to pardon one of the threes and chooses at random the prisoner to pardon. He informs the warden of his choice but requests that the name to be kept in secret for a few days.
  The next day, A tries to get the warden to tell him who had been pardoned. The warden refuses. A then asks which of B or C will be executed. The warden thinks for a while, then tells A that B is to be executed.
  What is the chance that A being pardoned given the information?

# Probability

- $P(\text{The warden tells that B gets executed} \mid \text{A is pardoned}) = 1/2$.
- $P(\text{The warden tells that B gets executed} \mid \text{C is pardoned}) = 1$
- $P(\text{The warden tells that B gets executed }) = 1/2*1/3 + 1*1/3 = 1/2$.
- $P(\text{ The warden tells that B gets executed and A is pardoned}) = 1/2*1/3=1/6$
- $P(\text{A is pardoned} \mid \text{The warden tells that B gets executed}) = 1/3$.

Watch the following video

`https://www.youtube.com/watch?v=4Lb-6rxZxx0`

# Probability

- Two events, $A$ and $B$, are statistically independent if $P(A \cap B) = P(A)P(B)$.
- A random variable is a function from a sample space $S$ into the real number.
- Cumulative distribution function $F_X(x) = P(X \leq x)$ for all $x$.
- Theorem
  $F(x)$ is a cdf if and only if
  1. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.
  2. $F(x)$ is a nondecreasing function of $x$.
  3. $F(x)$ is right continuous.

# Distributions of functions of a random variable

- (Theorem) $X \sim f_X(x)$
  $Y := g(X)$.
  Suppose there exists a partition, $A_0, A_1, \ldots, A_k$ of $\chi$ such that $P(X \in A_0) = 0$ and $f_X$ is continuous on each $A_i$. Further, suppose there exist functions $g_1(x), \ldots, g_k(x)$, defined on $A_1, \ldots, A_k$, respectively, satisfying

  1. $g(x) = g_i(x)$ for $x \in A_i$
  2. $g_i(x)$ is monotone on $A_i$
  3. $\mathscr{Y} = \{y : y = g_i(x) \text{ for some } x \in A_i\}$
  4. $g_i^{-1}(y)$ has a continuous derivative on $\mathscr{Y}$ for each $i = 1, \ldots, k$

  Then $f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right| & y \in \mathscr{Y} \\ 0 & \text{otherwise} \end{cases}$

- (Example) $X$ is a continuous random variable with pdf $f_X(x)$. Find pdf of $Y = X^2$.

- $A_1 = \{x : x < 0\}$ and $A_2 = \{x : x > 0\}$
- $\mathscr{Y} = \{y : y > 0\}$
- $g_1^{-1}(y) = -\sqrt{y}$, $g_2^{-1}(y) = \sqrt{y}$
- $\left|\frac{d}{dy} g_i^{-1}(y)\right| = \frac{1}{2\sqrt{y}}$
- $f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(-\sqrt{y}) + f_X(\sqrt{y}))$

# Distributions of functions of a random variable

- (Example) Let $X$ have the standard normal distribution, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$, $-\infty < x < \infty$.
  Consider $Y = X^2$. Find the pdf of $Y$.
  $\langle$*Solution*$\rangle$
  $f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(-\sqrt{y}) + f_X(\sqrt{y})) = \frac{1}{\sqrt{2\pi}} \exp(-y/2), 0 < y < \infty$
- (Theorem) Let $X$ have a continuous cdf $F_X(x)$ and define the random variable $Y$ as $Y = F_X(X)$. Then $Y$ is uniformly distributed on $(0,1)$, that is, $P(Y \le y) = y, 0 < y < 1$.
  $\langle$*Proof*$\rangle$
  $P(F_X(X) \le y) = P(X \le F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y$

# Moment generating function

- $M_X(t) = E(e^{tX})$, $-h < t < h$ and $h > 0$.
- (Theorem) $X$ has mgf of $M_X(t)$
  $EX^n = M_X^{(n)}(0)$,
  where $M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t)|_{t=0}$
- (Theorem) $F_X(x)$ and $F_Y(y)$: two cdfs all whose moments exist.
  1. If $X$ and $Y$ have bounded support,
     $F_X(u) = F_Y(u)$ for all $u \Leftrightarrow EX^r = EY^r$
  2. If the moment generating function exist
     $M_X(t) = M_Y(t)$ for all $t$ in some neighborhood of 0, then
     $F_X(u) = F_Y(u)$ for all $u$.

# Moment generating function

- (Theorem) $\{X_i, i = 1, 2, \ldots\}$ sequence of r.v.
  $\lim_{i\to\infty} M_{X_i}(t) = M_X(t)$ for all $t$ in a neighborhood of zero and $M_X(t)$ is an mgf.
  Then there is a unique cdf $F_X$ whose moments are determined by $M_X(t)$ and for all $x$ where $F_X(x)$ is conti
  $\lim_{i\to\infty} F_{X_i}(x) = F_X(x)$.

# Some useful theorems

## Theorem 2.4.2 (Dominated Convergence Theorem)

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

## Theorem 2.4.3

Suppose $f(x,\theta)$ is differentiable at $\theta = \theta_0$
$\lim_{\delta\to 0} \frac{f(x,\theta_0+\delta)-f(x,\theta_0)}{\delta} = \frac{\partial}{\partial\theta} f(x,\theta)|_{\theta=\theta_0}$ exists for every $x$.
There exists a function $g(x,\theta_0)$ and a constant $\delta_0$ such that

- $|\frac{f(x,\theta_0+\delta)-f(x,\theta_0)}{\delta}| \leq g(x,\theta_0)$ for all $x$ and $|\delta| < \delta_0$,
- $\int_{-\infty}^{\infty} g(x,\theta_0)dx < \infty$

Then $\frac{d}{d\theta} \int_{-\infty}^{\infty} f(x,\theta)dx|_{\theta=\theta_0} = \int_{-\infty}^{\infty} \frac{\partial}{\partial\theta} f(x,\theta)|_{\theta=\theta_0} dx$

# Some useful theorems

## Corollary 2.4.4

Suppose $f(x,\theta)$ is differentiable in $\theta$ and there exists a function $g(x,\theta)$ such that $|\frac{\partial}{\partial\theta}f(x,\theta)|_{\theta=\theta'}| \leq g(x,\theta)$ for all $\theta'$ such that $|\theta'-\theta| \leq \delta_0$ and $\int_{-\infty}^{\infty} g(x,\theta)dx < \infty$
then $\frac{d}{d\theta}\int_{-\infty}^{\infty} f(x,\theta)dx|_{\theta=\theta_0} = \int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}f(x,\theta)|_{\theta=\theta_0}dx$

- Example 3.4.5 $f_X(x) = \frac{1}{\lambda}e^{-x/\lambda}$
  Calculate $\frac{d}{d\lambda}EX^n = \frac{d}{d\lambda}\int_0^\infty x^n \frac{1}{\lambda}e^{-x/\lambda}dx$.

$$\begin{aligned} & \int_0^\infty \frac{\partial}{\partial\lambda}(x^n\frac{1}{\lambda}e^{-x/\lambda})dx \\ = & \int_0^\infty \frac{x^n}{\lambda^2}\left(\frac{x}{\lambda}-1\right)e^{-x/\lambda}dx \\ = & \frac{1}{\lambda^2}EX^{n+1} - \frac{1}{\lambda}EX^n. \end{aligned}$$

# Some useful theorems

- $\frac{x^n}{\lambda^2}\left(\frac{x}{\lambda}-1\right)e^{-x/\lambda} \leq \frac{x^n}{(\lambda-\delta_0)^2}\left(\frac{x}{(\lambda-\delta_0)}-1\right)e^{-x/(\lambda+\delta_0)}$ for some $0<\delta_0<\lambda$.
- Take $g(x,\lambda)=\frac{x^n}{(\lambda-\delta_0)^2}\left(\frac{x}{(\lambda-\delta_0)}-1\right)e^{-x/(\lambda+\delta_0)}$
- $\left|\frac{\partial}{\partial\lambda}(x^n\frac{1}{\lambda}e^{-x/\lambda})|_{\lambda=\lambda'}\right| \leq g(x,\lambda)$ for $|\lambda-\lambda'|<\delta_0$