Midterm 2 Study Guide

William Sun

Random Variables

- $X : \Omega \to \mathbb{R}$ can be discrete or continuous
- If X is a discrete random variable, we will associate it with a *probability mass function (pmf)* P_X
  - $P_X(x) > 0 \ \forall \ x \in \{\text{values of } X\}$
  - $\sum_x P_X(x) = 1$, where the sum is over all the possible values of X
  - Used to calculate some probabilities: $P(X \in A) = \sum_{x \in A} P_X(x)$
- If X is a continuous random variable, we will associate it with a *probability density function (pdf)* of f
  - $f : \mathbb{R} \to \mathbb{R}, f(x) \geq 0 \forall x \in \mathbb{R}$
  - $\int\limits_{-\infty}^{\infty} f(x)dx = 1$

Functions of Random variables

- With pmf of X and $Y = g(X)$:

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x)$$

- If X is a continuous random variable with pdf $f_X(x)$ and $Y = g(X)$ where g is a monotone (increasing or decreasing) then the pdf of Y is given by

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{d}{dy}(g^{-1}(y)) \right|$$

  - Or since $x = g^{-1}(y)$,

$$f_Y(y) = f_X(x)\frac{dx}{dy}.$$

- Suppose $X \sim f_X(x)$ and $Y = g(X)$. What is the pdf of Y?
  - Step 1: Compute the cdf of Y in terms of the cdf of X

- Step 2: Take a derivative and use the chain rule
- Example in March 27th notes

Expected Value and Variance of Random Variable

- $\mathbb{E}$: also the mean, weighted average, or center of mass
- Discrete:

$$\mathbb{E}(X) = \sum_x x P_X(x)$$

- Continuous:

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

- The variance of an random variable: $Var(x) = (x - \mu)^2$, where $\mu = \mathbb{E}(X)$
  - So $\mathbb{E}(\{X - \mu\}^2) := Var(X)$
- A form of the $Var(X)$ more amenable to calculations: $\mathbf{Var(X)} = \mathbb{E}(\mathbf{X^2}) - \mathbb{E}(\mathbf{X})^\mathbf{2}$

Cumulative Distribution Function (CDF)

- $F(X) = P(X \leq x)$
  1) $F : \mathbb{R} \to [0, 1]$
  2) If $x < y$, $F(x) \leq F(y)$
- Notation: Left-limit notation $F(c-) = lim_{x \to c-} F(x)$
- For continuous random variables

$$P(-\infty < X \leq x) = F_X(x) = \int_{-\infty}^{x} f(u) du$$

- If we know the CDF,
  - $P(a < x \leq b) = F(b) - F(a)$
  - $P(a \leq x \leq b) = F(b) - F(a-)$
  - $P(a \leq x < b) = F(b-) - F(a-)$
  - $P(a < x < b) = F(b-) - F(a)$
  - General rule: If near "<", $a \to -F(a)$, $b \to F(b-)$
    - "a < b" ... so if "<" is near a, the "-" is on the left; "-" is on the right for b

Law of the Unconscious Statistician

- If X is discrete and $G : \mathbb{R} \to \mathbb{R}$, then
  - $\mathbb{E}(\mathbf{g}(\mathbf{X})) = \sum_{\mathbf{x}} \mathbf{g}(\mathbf{x})\mathbf{P}(\mathbf{X} = \mathbf{x})$ when the expectation exists
- Continuous:
  - If X is a continuous random variable with pdf $f(x)$ and $g : \mathbb{R} \to \mathbb{R}$ is any function such that

$$\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty$$

(This is a condition which will guarantee that the expected value exists and is finite.) then

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

- Used when we know $G(X)$ and the distribution of X but not the distribution of $G(X)$

Linearity of Expectation

- Linearity of Expectation #1
  - $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
- Linearity of Expectation #2
  - $\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}X_n)$
  - Expectation of a sum is the sum of the individual expected values
  - For any random variables for which $\mathbb{E}(X_i)$ exists for all i

Normal Random Variables

- Theorem: If $X \sim \text{Normal}(\mu, \sigma^2)$ and a, b are any constants, then
  - $Y = aX + b \sim \text{Normal}(a\mu + b, a^2\sigma^2)$
  - If you have a normal random variable, any linear transformation on it is also a normal random variable.
  - *Consequence:* $X \sim \text{Normal}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \text{Normal}(0,\, 1)$

- * Any normal random variable can be converted into a *standard normal distribution* with a mean of 0 and a standard deviation of 1

- To compute the probability of a normal random variable, use the z-table.

The Euler Gamma Function

- $$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}dy$$
- $\Gamma(a+1) = a\Gamma(a) \to \Gamma(a) = (a-1)!$ for $a > 0$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

The Normalization Trick

- Remark: By recognizing this pdf in one form another and using the fact
- $$\int_0^\infty x^{\alpha-1}e^{-\frac{x}{\beta}}dx = \beta^\alpha\Gamma(\alpha)$$
- Will allow us to compute $\mathbb{E}(X^n)$ $\forall$ n.
  - Example in March 27 notes

Weibull distribution

- $X \sim \exp(1)$ $f_X(x) = e^{-x}$ for $x > 0$
- Find pdf of $Y = \nu + \alpha X^{\frac{1}{\beta}}$ $(\nu \in \mathbb{R}, \alpha > 0, \beta > 0)$

Joint Distribution

- When X, Y are jointly discrete, we define the *joint pmf*

$$P_{X,Y}(x,y) := P(X = x, Y = y)$$

- Which is shorthand for $P(\{X = x\} \cap \{Y = y\})$

Marginal PDFs

- The marginal pdf of X
  - $f_X(x) = \int\limits_{-\infty}^{\infty} f(x, y)dy$
- and the marginal pdf of Y
  - $f_Y(y) = \int\limits_{-\infty}^{\infty} f(x, y)dx$
- If function is $f(x_1, x_2, x_3, x_4, x_5)$, $f_{x_1}(x_1) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x_1, x_2, x_3, x_4, x_5)dx_2dx_3dx_4dx_5$
  - Also extends to multivariate e.g. $f_{x_2,x_4}(x_2, x_4)$
  - Basically, integrate out everything that's not the thing you're interested in.

Independence

- $X_1, X_2, \ldots X_n$ jointly distributed random variables
- We'll say they are independent if joint distribution $= \prod\limits_{i=1}^{n}$ marginal distribution
  - $p(x_1, x_2, \ldots x_n) = P_{X_1}(x_1)P_{X_2}(x_2) \ldots P_{X_n}(x_n) \; \forall \; x_1, x_2, \ldots x_n$

Sums of Random Variables

- If X_1, X_2 are jointly distributed random variables, then what is the distribution (pmf) of $X_1 + X_2$?
- Case 1: Suppose X_1, X_2 jointly discrete, $P_{X_1,X_2}(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$ given
  - $P_{X_1,X_2}(u) = P(X_1 + X_2 = u) = \sum\limits_{x_1} P(X_1 + X_2, X_1 = x_1) = \sum\limits_{x_1} P(X_1 = x_1, X_2 = u - x) = \sum\limits_{x_1} P_{X_1,X_2}(x_1, u - x_1)$
    - (Law of total probability)
  - Formula:

$$P_{X_1+X_2}(u) = \sum_{x_1} P_{X_1,X_2}(x_1, u - x_1)$$

  - A common assumption is that X_1, X_2 independent. In this case:
    - $$P_{X_1+X_2}(u) = \sum_{x_1} P_{X_1}(x_1)P_{X_2}(u - x_1)$$

    - $*$ Convolution $(P_{X_1} * P_{X_2})(u)$
- Binomial theorem $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$
- X, Y jointly continuous with joint pdf $F(x, y)$. PDF of $X + Y$ is

$$\int_{-\infty}^{\infty} f(u-y, y)dy \text{ or } \int_{-\infty}^{\infty} f(x, u-x)$$

- If X, Y are independent
  -
$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f_X(x) f_Y(u-x)dx$$
    - $*$ *Convolution integral*
  - If $X \geq 0, Y \geq 0$ then
$$f_{X+Y}(u) = \int_{0}^{u} f_X(x) f_Y(u-x)dx$$

Ordered Statistics

- $X_1, X_2 \ldots X_n \sim$ independent, continuous random variables all having the same distribution (iidf)
  - $X_{(1)}$ = smallest among $X_1, X_2 \ldots X_n$
  - $\vdots$
  - $X_{(j)}$ = jth smallest among $X_1, X_2 \ldots X_n$
  - $\vdots$
  - $X_{(n)}$ is the largest
-

Distributions:

$$f_{Y_1}(y) = n(1-F(y))^{n-1} f(y)$$
$$f_{Y_j}(y) = \frac{n!}{(j-1)!(n-j)!} F(y)^{j-1} f(y)(1-F(y))^{n-j}$$
$$f_{Y_n}(y) = n(F(y))^{n-1} \cdot f(y)$$

Conditional Distributions

- Suppose X, Y are jointly continuous
- Define $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$ assuming $f_X(x) > 0$.
- Application to sum of random variables:

$$f_{X|X+Y}(x|u) = \frac{f_{X,Y}(x, u-x)}{f_{X+Y}(u)}$$

Transformation Theorem (Method of Jacobians)

- For finding distributions of functions of continuous random variables
- (2-d) Theorem: Suppose X, Y are jointly continuous with joint pdf $f_{X, Y}(x, y) with support A and $u = g_1(x, y)$ and $v = g_2(x, y)$ is a one-to-one transformation of A into B. Then the inverse transformation is

$$x = h_1(u, v) \text{ and } y = h_2(u, v)$$

- and the joint pdf of U, V is of the form

$$f_{U,V}(u, v) = f_{X,Y}(h_1(u, v), h_2(u, v))|J|$$

- where J is the determinant of $\begin{bmatrix} \frac{dx}{du} & \frac{dx}{dv} \\ \frac{dy}{du} & \frac{dy}{dv} \end{bmatrix} = \frac{dx}{du} \cdot \frac{dy}{dv} - \frac{dx}{dv} \cdot \frac{dy}{du}$ (the Jacobian determinant)

Bivariate Normal

- Bivariate Normal X, Y with
  - $\mu_X, \mu_Y \in \mathbb{R}$ means
  - σ_X, σ_Y 0 standard deviations
  - $-1 < \rho < 1$ correlation coefficient
  - Let Z_1, Z_2 be independent standard normal

$$f_{Z_1,Z_2}(z_1, z_2) = \frac{exp(-\frac{z_1^2+z_2^2}{2})}{2\pi}$$

  - $Y|X$ gets complicated, but we can simplify it to $Y|X = x \sim \text{Normal}(\mu_Y + \sigma_Y \rho(\frac{x-\mu_X}{\sigma_X}, \sigma_Y^2(1-\rho^2))$

Fisher's F-distribution

- $X \sim \chi^2_n, Y \sim \chi^2_m$ and they are independent
- Then define

$$U = \frac{X/n}{Y/m}, V = Y$$

- For $u > 0$, (show as an exercise)

$$f_U(u) = \frac{n\Gamma(\frac{n+m}{2})}{m\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} \cdot \frac{(\frac{nu}{m})^{\frac{n}{2}-1}}{(1+\frac{nu}{m})^{\frac{n+m}{2}}}$$

  - The F-distribution with n Numerator d.f. (degrees of freedom), m Denominator d.f.
  - Remember: If $W \sim$ Cauchy, $W^2 = F_{1, 1}

Extra Credit: An Identity About the Beta-Family of PDFs

- $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$