What to study today (Sept 28, 2020)?

2. Probability and Distribution (Chp 1-3)

2.1 Probability (Chp1.1-4)

2.2 Random Variable and Distribution (Chp1.5-10)

- *2.2.1 Basic Concepts*
- *2.2.2 Discrete Random Variable*
- *2.2.3 Continuous Random Variable*
- **2.2.4 Expectation and Related**

2.3 Multivariate Distribution (Chp2)

- **2.3.1 Basic Concepts with Two Random Variables**
- *2.3.2 Conditional Distribution and Expectation*
- *2.3.3 Extension to Several Random Variable*

2.4 Some Special Distributions (Chp3)

Example 2.6 Assume r.v. $W \sim U(0,1)$, the continuous uniform distribution over $[0,1]$.

(1) Define a r.v. $X = aW + b$. What are the pdf and cdf of X?

(2) Define another discrete r.v. $Y = \begin{cases} 0, & 0 \leq W < 1/3 \\ 1, & 1/3 \leq W \leq 1 \end{cases}$.
What are the pmf and cdf of Y?

Review: Expectation

Definition. The **expectation** of rv X with cdf $F(x)$ is $E(X) = \int_{-\infty}^{\infty} x dF(x)$, provided the integral exists.

- If X is continuous with pdf $f(x)$, $dF(x) = f(x)dx$ and $E(X) = \int_{-\infty}^{\infty} xf(x)dx$.
- If X is discrete with pmf $p(x)$, $dF(x) = F(x) - F(x-) = p(x)$ and $E(X) = \sum_{all\ x} xp(x)$.

Definition. The **expectation** of $g(X)$ with $X \sim F(x)$ is $E\big[g(X)\big] = \int_{-\infty}^{\infty} g(x)dF(x)$, provided the integral exists.

- The **variance** of rv X with cdf $F(x)$ and $\mu = E(X)$ is $Var(X) = E\big[(X-\mu)^2\big]$, provided the expectation exists.
- The **moment generating function** (mgf) of rv X with X's cdf $F(x)$ is $M(t) = E(e^{tX})$ for $t \in (-h, h)$, provided the expectation exists.
  - $M(0) = 1$; $M^{'}(0) = \frac{dM(t)}{dt}\big|_{t=0} = E(X)$; $M^{''}(0) = E(X^2)$; $M^k(0) = \frac{d^k M(t)}{dt^k}\big|_{t=0} = E(X^k)$ with any integer k.
  - Two rvs X and Y have the same mgf iff their cdf are the same.

2.2.4 Expectation and Related: Further Properties

- The expectation operator E is linear: for constants k_1, k_2,

$$E\left[k_1 g_1(X) + k_2 g_2(X)\right] = k_1 E\left[g_1(X)\right] + k_2 E\left[g_2(X)\right]$$

provided the expectations exist.

For example, if the pdf of rv X is $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ for $-\infty < x < \infty$, what is $E(3X^2 - 2X - 1)$? [2]

- $Var(k_1 X + k_2) = k_1^2 Var(X)$.

For example, if the pdf of rv X is $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ for $-\infty < x < \infty$, what is $Var(3X - 2)$? [9]

Example 2.7 Assume r.v. $W \sim U(0,1)$, the continuous uniform distribution over $[0,1]$.

(1) Define a r.v. $X = aW + b$. What are $E(X)$ and $Var(X)$? What is the mgf of X?

(2) Define another discrete r.v. $Y = \begin{cases} 0, & 0 \leq W < 1/3 \\ 1, & 1/3 \leq W \leq 1 \end{cases}$. What are $E(Y)$ and $Var(Y)$? What is the mgf of Y?

2.2.4 Expectation and Related: Further Properties

- **Markov's Inequality.** If $h(x)$ is a nonnegative function, for $a > 0$, $P\left[h(X) \geq a\right] \leq \frac{E\left[h(X)\right]}{a}$.
  - A special case: **Chebyshev's Inequality.** for constant k, $P\left[|X - \mu| \geq k\sigma\right] \leq \frac{1}{k^2}$ with $\mu = E(X)$ and $\sigma^2 = Var(X)$.

- **Jensen's Inequality.** If $\phi(\cdot)$ is convex (convex downward or concave upward) on rv X's support, $\phi\left[E(X)\right] \leq E\left[\phi(X)\right]$.
  - A special case: $\left[E(X)\right]^2 \leq E(X^2)$.

2.3 Multivariate Distribution (Chp2): Basic Concepts

Consider two rvs X_1, X_2: study them one at a time via the tools discussed in Chp 1?

Definition. The **joint cdf** of the random vector (X_1, X_2) is

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$$

for $-\infty < x_1, x_2 < \infty$. If the sample space is S,
$F(x_1, x_2) = P\Big(\{c : X_1(c) \leq x_1\} \bigcap \{c : X_2(c) \leq x_2\}\Big)$.

Basic Properties of $F(x_1, x_2)$:

- $0 \leq F(x_1, x_2) \leq 1$ for all $-\infty < x_1, x_2 < \infty$.
- nondecreasing and right-continuous wrt either of x_1, x_2.
  - $\lim_{x_1 \to \infty, x_2 \to \infty} F(x_1, x_2) = 1$.
  - $\lim_{x_1 \to -\infty} F(x_1, x_2) = \lim_{x_2 \to -\infty} F(x_1, x_2) = 0$
  - $\lim_{x_1 \to \infty} F(x_1, x_2) = P(X_2 \leq x_2) = F_{X_2}(x_2)$, the **marginal** cdf of X_2; $\lim_{x_2 \to \infty} F(x_1, x_2) = P(X_1 \leq x_1) = F_{X_1}(x_1)$, the **marginal** cdf of X_1

2.3.1 Basic Concepts with Two Random Variables

- When X_1, X_2 are discrete, the **joint pmf** of (X_1, X_2) is $p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$.

Example 2.8 Consider to roll a die after flipping an even coin: use a six-face die with face numbers 1,1,2,2,3,3 if getting a head; otherwise, use a die with numbers 1,1,1,2,2,2. Let X_1 be the number of head and X_2 be the number from the die.

- What is the joint pmf of (X_1, X_2)?
- What is the pmf of X_1, of X_2?
- In general, the (**marginal**) pmf of X_1 is $p_{X_1}(x_1) = \sum_{all\ x_2} p(x_1, x_2)$; the pmf of X_2 is $p_{X_2}(x_2) = \sum_{all\ x_1} p(x_1, x_2)$
- Can we obtain the joint pmf $p(x_1, x_2)$ from the marginal pmf $p_{X_1}(\cdot)$ and the marginal pmf $p_{X_2}(\cdot)$?

- When X_1, X_2 are continuous, the **joint pdf** of (X_1, X_2) is $f(x_1, x_2)$ such that $P\big((X_1, X_2) \in A\big) = \int\int_A f(x_1, x_2) dx_1 dx_2$.

Example 2.9 Consider two continuous rvs X_1, X_2 with joint pdf:

$$f(x_1, x_2) = \begin{cases} x_1 + x_2, & 0 < x_1, x_2 < 1 \\ 0, & otherwise \end{cases}$$

- What is $P(X_1 \geq X_2) =$? [1/2] What is $P(X_1 + X_2 \leq 1) =$? [1/3]

- What is the (marginal) pdf of X_1, of X_2?

- Can we obtain the joint pdf $f(x_1, x_2)$ from the marginal pdf $f_{X_1}(\cdot)$ and the marginal pdf $f_{X_2}(\cdot)$?

2.3.1 Basic Concepts with Two Random Variables

Definition. Two rvs X_1, X_2 are **independent** iff their joint cdf is the product of their marginal cdfs: $F(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2)$ for $-\infty < x_1, x_2 < \infty$, that is
$P(X_1 \leq x_1, X_2 \leq x_2) = P(X_1 \leq x_1)P(X_2 \leq x_2)$ for $-\infty < x_1, x_2 < \infty$,

- Two discrete rvs X_1, X_2 are **independent** iff their joint pmf is the product of their marginal pmfs: $p(x_1, x_2) = p_{X_1}(x_1)p_{X_2}(x_2)$.

- Two continuous rvs X_1, X_2 are **independent** iff their joint pdf is the product of their marginal pdfs: $f(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$.

2.3.1 Basic Concepts with Two Random Variables

Definition. The expectation of $Y = g(X_1, X_2)$, a function of two rvs $X_, X_2$, is $E(Y) = \int\int g(x_1, x_2) dF(x_1, x_2)$.

- If X_1, X_2 are both discrete with joint pmf $p(x_1, x_2)$, $E(Y) = \sum_{all\ x_1, x_2} g(x_1, x_2) p(x_1, x_2)$.
- If X_1, X_2 are both continuous with joint pdf $f(x_1, x_2)$, $E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) f(x_1, x_2) dx_1 dx_2$.

Definition. The **covariance** of two rvs X_1, X_2 is $Cov(X_1, X_2) = E\big[(X_1 - \mu_1)(X_2 - \mu_2)\big] = E(X_1 X_2) - \mu_1 \mu_2$, where $\mu_1 = E(X_1)$ and $\mu_2 = E(X_2)$.

- $Cov(X_1, X_1) = Var(X_1)$
- The **correlation coefficient** of X_1, X_2 is $\rho(X_1, X_2) = Cov(X_1, X_2) \big/ \sqrt{Var(X_1) Var(X_2)}$: $-1 \le \rho \le 1$.

2.3.1 Basic Concepts with Two Random Variables

- If X_1, X_2 are indpt ($X_1 \perp\!\!\!\perp X_2$), $E(X_1 X_2) = E(X_1)E(X_2)$, and thus $Cov(X_1, X_2) = 0$ and $\rho(X_1, X_2) = 0$.

- **Example 2.8** (cont'd) What are $Cov(X_1, X_2)$, $\rho(X_1, X_2)$?

- **Example 2.9** (cont'd) What are $Cov(X_1, X_2)$, $\rho(X_1, X_2)$?

What will we study in the next class?

1. *Introduction*
2. **Probability and Distribution (Chp 1-3)**
   - *2.1 Probability (Chp1.1-4)*
   - *2.2 Random Variable and Distribution (Chp1.5-10)*
   - **2.3 Multivariate Distribution (Chp2)**
     - *2.3.1 Basic Concepts with Two Random Variables*
     - **2.3.2 Conditional Distribution and Expectation**
     - **2.3.3 Extension to Several Random Variables**
   - *2.4 Some Important Distributions (Chp3)*
3. *Essential Topics in Mathematical Statistics (Chp 4-6)*
4. *Further Topics, Selected from Chp 7-11*