Chapter 2: Random Variables and their Distributions

Model outcomes in world with random vars, which give probs that the variable takes on a given value. Discrete and Continuous (continuous often easier as have integration theory).

Section 2.1: Random Variables

Defn: A random variable is a function $X: \Omega \to \mathbb{R}$ with the property that $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for each $x \in \mathbb{R}$; such a function is called $\mathcal{F}$-measurable

Example: Toss a fair coin twice, so $\Omega = \{HH, HT, TH, TT\}$. For $\omega \in \Omega$, let $X_0(\omega) = \#$ heads and $X_2(\omega) = \#$ heads in second toss, ~~thus~~ $X_1 = \#$ heads in first toss. Then

$X_0(HH) = 2,\ X_0(HT) = X_0(TH) = 1,\ X_0(TT) = 0$

$X_2(HH) = X_2(TH) = 1, \quad X_2(HT) = X_2(TT) = 0$

$X_1(HH) = X_1(HT) = 1, \quad X_1(TH) = X_1(TT) = 0$

↳ Note $X_0 = X_1 + X_2$

(*) IMPORTANT! WHY DO WE WANT $X(\omega)$ real valued?

↳ answer: This way can add (apples and apples vs oranges and apples).

↳ ex: Say $Y_1(\omega) = H$ if first toss is a head and a tail otherwise and $Y_2(\omega) = H$ " second " " " " " " " "

What is $Y_1(\omega) + Y_2(\omega)$? What is a head plus a tail?

Notation: UPPER CASE $X, Y, Z, \ldots$ denote random variables

lower case $x, y, z, \ldots$ denote values of the random vars

SECTION 2.1: (CONTINUED) RANDOM VARIABLES

- discrete random vars: $f(x)$ is prob random variable X equals x.
- Continuous random vars: careful, as singletons have zero probability

Defn: The distribution fn of a random variable $X : \Omega \to \mathbb{R}$ is the function $F : \mathbb{R} \to [0,1]$ given by $F(x) = \mathbb{P}(X \le x)$

Notation: See calculus notation: F is anti-deriv of f

Say X continuous: $\mathbb{P}(X \le x) = \int_{-\infty}^{x} f(t)\,dt = F(x)$

↳ by Fund Thm of Calc, F is continuous for f "nice"

↳ Sometimes write F_X to remind ourselves what is the rand var

Example: Coin Problem: $F_{X_0}(x) = \begin{cases} 0 & x < 0 \\ 1/4 & 0 \le x < 1 \\ 3/4 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$ $\quad F_{X_1} = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$

LEMMA: F is a distribution function if and only if it satisfies:

(1) $\lim_{x \to -\infty} F(x) = 0 \qquad \lim_{x \to \infty} F(x) = 1$

(2) If $x < y$ then $F(x) \le F(y)$

(3) F is right continuous: $\lim_{h \to 0^+} F(x+h) = F(x)$

Proof of (3): Lemma 1.3.5 (pg 7): events $B_1 \supset B_2 \supset \cdots$ Then

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} B_i\right) = \lim_{i \to \infty} \mathbb{P}(B_i)$$

Let $B_i = \{\omega \in \Omega : X(\omega) \le x + 1/i\}$

↳ Note $\bigcap_{i=1}^{\infty} B_i = \{\omega \in \Omega : X(\omega) \le x\}$

Section 2.1: Random Vars and Distr (cont)

Examples

- Constant: $X(\omega) = c$ for all $\omega \in \Omega$
- Constant (almost surely): $\mathbb{P}(X = c) = 1$

  ↳ distributive fn is $F_X(x) = \begin{cases} 0 & x < c \\ 1 & x \geq c \end{cases}$
- Bernoulli (Bern(p)): $\Omega = \{H, T\}$, $\mathbb{P}(H) = p$, $\mathbb{P}(T) = 1-p$

$$X(H) = 1, \quad X(T) = 0$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

  ↳ often write $X \sim \text{Bern}(p)$
- Indicator fns: Event $A \subset \Omega$, $I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$

  ↳ note I_A is $\text{Bern}(\mathbb{P}(A))$

  ↳ often write $\mathbb{1}_A$ or χ_A

  ↳ useful identity: $A \subset \bigcup_{i \in I} B_i$ and B_i's distinct, $I_A = \sum_{i \in I} I_{A \cap B_i}$

Lemma: F distribution fn for X then

(1) $\mathbb{P}(X > x) = 1 - F(x)$

(2) $\mathbb{P}(x < X \leq y) = F(y) - F(x)$

(3) $\mathbb{P}(X = x) = F(x) - \lim_{y \to x^-} F(y)$

Proof: for (3) use $B_n = \{x - 1/n < X \leq x\}$

Notation: Tails $T_1(x) = \mathbb{P}(X > x) = 1 - F(x)$

$$T_2(x) = \mathbb{P}(X \leq -x) = F(-x)$$

HW do: #2, #4, #5c

suggested: #1, #3, #5abd

Section 2.2: The Law of Averages

↳ Optional reading: will deal with such convergence questions in greater detail later in the semester

Section 2.3: Discrete and Continuous Variables

Defns • Random variable X is discrete if takes values in a countable subset $\{x_1, x_2, \dots\}$ of $\mathbb{R}$; has (probability) mass function $f: \mathbb{R} \to [0,1]$ given by $f(x) = \mathbb{P}(X = x)$

• Random variable X is continuous if its distribution function can be written as $F(x) = \int_{-\infty}^{x} f(u)\,du \quad (x \in \mathbb{R})$ for some integrable function $f: \mathbb{R} \to [0,\infty)$ called the (probability) density fn of X.

↳ Notes • countable means either finite or $\exists$ 1-1 and onto function from our set to $\mathbb{N}$ (some books call this at most countable) Give examples

• integrable: think Riemann-integrable

↳ Not all random vars of this form; more advanced analysis (Lebesgue-Stieltjes integrals) allow us to consider large class of random variables

Example: Discrete: $X_i : \{H,T\}^2 \to \mathbb{R}$ by $X_i(\omega) = \begin{cases} 1 & i^{th} \text{ toss } H \\ 0 & i^{th} \text{ toss } T \end{cases}$

Continuous: $X \sim Exp(\lambda)$: $F(x) = \int_{0}^{x} \frac{1}{\lambda} e^{-x/\lambda}\,dx$

Example: Change of Variable: Say $X \sim Exp(\lambda)$: What is $Y = X^2$

↳ Use distribution fns (also called cumulative distribution fn)

Let $G(y) = \mathbb{P}(Y \le y) = \mathbb{P}(X^2 \le y) = \mathbb{P}(X \le \sqrt{y}) = F_X(\sqrt{y})$

$$= \int_{0}^{\sqrt{y}} \frac{1}{\lambda} e^{-x/\lambda}\,dx = \int_{0}^{\sqrt{y}/\lambda} e^{-u}\,du = 1 - e^{-\sqrt{y}/\lambda}$$

so $g(y) = G'(y) = \frac{1}{2}\frac{1}{\sqrt{y}}\frac{1}{\lambda} e^{-\sqrt{y}/\lambda}$

SECTION 2.3: DISCRETE + CONT RAND VARS (CONT)

Example: Faster soln

$$G(Y) = F_X(\sqrt{Y})$$

↳ Thus $g(Y) = G'(Y) = F_X'(\sqrt{Y}) \cdot (\sqrt{Y})'$ Chain Rule

$$= f_X(\sqrt{Y}) \cdot \frac{1}{2\sqrt{Y}}$$

$$= \frac{1}{2\sqrt{Y}} \frac{1}{\lambda} e^{-\sqrt{Y}/\lambda}$$

↳ note this method does not require us to know F_X, only f_X

↳ not surprising, as only need derivative (i.e. the density fn)

HW: do: #3*, #4, #5

Suggested: #1, #2

* Extremely important problem!

Allows us to use the distribution function of the uniform random variable on [0,1] to generate other random variables; in other words, if we can choose $X \sim Unif(0,1)$, we can choose Y having many nice distributions

Section 2.4: Worked Examples

↳ can read but won't discuss in class

Section 2.5: Random Vectors

<u>Defn</u>: Joint distributions function of a random vector $\vec{X} = (X_1, \dots, X_n)$ on Probability Space $(\Omega, \mathcal{F}, \mathbb{P})$ is the function $F_{\vec{X}} : \mathbb{R}^n \to [0,1]$ given by $F_{\vec{X}}(\vec{x}) = \mathbb{P}(\vec{X} \le \vec{x})$ for $\vec{x} \in \mathbb{R}^n$, where $\vec{x} \le \vec{y}$ means each $x_i \le y_i$ and $\{\vec{X} \le \vec{x}\} = \{\omega \in \Omega : \vec{X}(\omega) \le \vec{x}\}$

Lemma: Joint Distribution Function $F_{\vec{X}}$ of the random vector $\vec{X}$ satisfies (1) $\lim_{x_1, \dots, x_n \to -\infty} F_{\vec{X}}(\vec{x}) = 0$, $\lim_{x_1, \dots, x_n \to \infty} F_{\vec{X}}(\vec{x}) = 1$

(2) If $\vec{x} \le \vec{x}'$ then $F_{\vec{X}}(\vec{x}) \le F_{\vec{X}}(\vec{x}')$

(3) $F_{\vec{X}}$ cont from above: $\lim_{\vec{u} \to \vec{0}^+} F_{\vec{X}}(\vec{x} + \vec{u}) = F_{\vec{X}}(\vec{x})$

<u>Defn</u>
- $X_1, \dots, X_n$ random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, say <u>jointly discrete</u> if $\vec{X} = (X_1, \dots, X_n)$ takes values in a countable subset of $\mathbb{R}^n$, and have joint (probability) mass function $f : \mathbb{R}^n \to [0,1]$ given by $f(x_1, \dots, x_n) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$.
- Same setup, <u>jointly continuous</u> if $F_{\vec{X}}(\vec{x}) = \int_{u_1=-\infty}^{x_1} \cdots \int_{u_n=-\infty}^{x_n} f(u_1, \dots, u_n)\, du_1 \dots du_n$ for some integrable function $f : \mathbb{R}^n \to [0, \infty)$

<u>Marginals</u>: $\lim_{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n \to \infty} F_{\vec{X}}(\vec{x}) = F_{X_j}(x_j)$

↳ called the marginals of $F_{\vec{X}}$

generally <u>not</u> possible to recapture $F_{\vec{X}}$ from the F_{X_j}

Challenge: Find a set of marginals that can correspond to at least two different $F_{\vec{X}}$'s.

SECTION 2.4: RANDOM VECTORS

__Examples:__ dart hits circular disk of radius ρ, probability in any "nice" subset of disk is proportional to area of region.

↳ area of disk is $\pi\rho^2$

$$\mathbb{P}(R \le r) = r^2/\rho^2 \qquad R = \text{radius location}$$

$$\mathbb{P}(\Theta \le \theta) = \theta/2\pi \qquad \Theta = \text{angle from } y\text{-axis}$$

$$\mathbb{P}(R \le r, \Theta \le \theta) = \mathbb{P}(R \le r)\,\mathbb{P}(\Theta \le \theta)$$

$$= \int_{u=0}^{r} \int_{v=0}^{\theta} f(u,v)\,du\,dv$$

$$\hookrightarrow f(u,v) = \frac{u}{\pi\rho^2}$$

$$0 \le u \le \rho, \quad 0 \le v \le 2\pi$$

__Example: Coin tossing__: $X_i = \begin{cases} 1 & i^{th} \text{ toss } H \\ 0 & i^{th} \text{ toss } T \end{cases}$, $\vec{X} = (X_1, \ldots, X_n)$

↳ encodes result of experiment

__Example: 3 sided coin (Head, Tail, Edge)__

↳ toss n times, H_n = # heads, T_n = # tails, E_n = # edges, all equally likely

$$\mathbb{P}((H_n, T_n, E_n) = (h_n, t_n, e_n)) = \begin{cases} \dfrac{n!}{h!\,t!\,e!} \left(1/3\right)^n & h, t, e \text{ non-neg integers that sum to } n \\ 0 & \text{otherwise} \end{cases}$$

↳ Proof: Binomial twice:

↳ two outcomes: head (1/3) and not-head (2/3)

$$\mathbb{P}(H_n = h, T_n + E_n = n-h) = \frac{n!}{h!\,(n-h)!}\left(\frac{1}{3}\right)^h \left(\frac{2}{3}\right)^{n-h}$$

$$\mathbb{P}(T_n = t, E_n = e \mid T_n + E_n = n-h) = \frac{(n-h)!}{t!\,(n-h-t)!}\left(\frac{1}{2}\right)^t \left(\frac{1}{2}\right)^{n-h-t}$$

$\cancel{\mathbb{P}((H_n, T_n, E_n) = (h,t,e)) = \mathbb{P}(H_n = h}$

HW: do: #2, #6

Sugg: #1, #4

Section 2.6: Monte Carlo Simulation

↳ will return to example 3 after discuss moments of a distribution (in particular, first two moments: mean and variance). We'll prove Chebychev's inequality (see Section 7.3, pg 319).

↳ Monte Carlo integration techniques extremely important
  ↳ most integrals cannot be done in closed form!

Section 2.7: Problems

HW: do:	#1,	#4,	#7,	#11,	#18
Can do: after read	§2.1	§2.3	§2.1	§2.1	§2.3

Suggested: #2, #3, #8, #12, #13, #15, #20