Last Class

- Review of basic probability

  - Sample space Ω = 'data structure' you need for any possible sim event

  - Probability - Collection of numbers, 'events' one for each set $S \subseteq \Omega$

    $\mathbb{P}[S] \geq 0 \ \forall S$, $\mathbb{P}[\Omega] = 1$

    $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] \ \forall A, B$ (mutually exclusive events) s.t. $A \cap B = \phi$

  - For any r.v. $X \in \mathbb{R}$

CDF $F_X(x) = \mathbb{P}[X \leq x]$

- Any CDF $F(x)$ has 3 properties

  1) $\lim_{x \to -\infty} F(x) = 0$, 2) $\lim_{x \to \infty} F(x) = 1$

  3) $F(x)$ is non-decreasing

Also one convention - in case of discontinuities, upper circle is filled in

Plan for today

- Relate CDFs to pmf (discrete) / pdf (continuous)

- Define expectation | variance

  - explain why this is imp for simulation

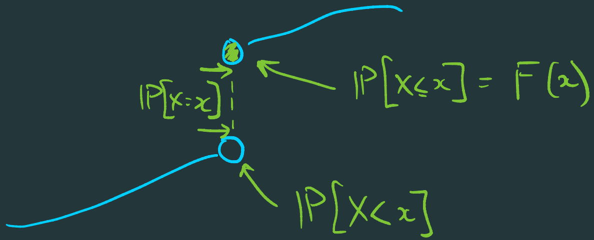
note on end-points

we wrote: $\mathbb{P}[a < X \le b] = F(b) - F(a)$: is $<$ vs $\le$ important?

because $$F(x) \triangleq \mathbb{P}[X \le x] = \underbrace{\mathbb{P}[X < x]} + \underbrace{\mathbb{P}[X = x]}$$

'is defined as'

both could be >0

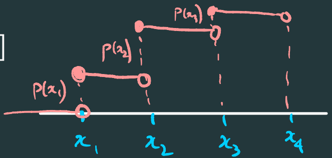


- for a discrete random variable, another characterization is its probability mass function (pmf) $p(\cdot)$

set Ω is the integers

cdf ≡ 'step function'

$$p(x) = \mathbb{P}[X = x]$$



- The pmf $p(\cdot)$ is related to the cdf $F(\cdot)$ as

$$F(x) = \sum_{x_i \leq x} p(x_i)$$

$$p(x) = F(x) - \max_{x_i < x} F(x_i) = F(x_i) - F(x_{i-1}) \text{ if } x = x_i$$

- further, any pmf $p(x)$ obeys 2 properties:

$$\cdot \quad \sum_{x_i} p(x) = 1 \quad , \quad p(x) \geq 0 \;\; \forall\, x \in \{x_1, x_2, \ldots\}$$

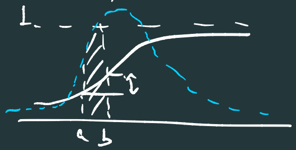
$$\Rightarrow \quad p(x) \leq 1 \;\; \forall\, x$$

continuous random variables

- for a continuous random variable taking values in $\mathbb{R}$, another characterization is its probability density function (pdf) $f(\cdot)$

$\Omega = \mathbb{R}$

cdf is a continuous fn



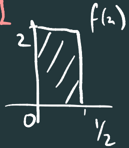
$$\mathbb{P}[a < X \leq b] = F(b) - F(a)$$
$$= \int_a^b f(x)\,dx$$

- any pdf $f(x)$ obeys 2 properties:

1) $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Note - $f(x)$ can be > 1

Eg - Unif$[0, 1/2]$



- ALERT!! not true that $f(x) = \mathbb{P}[X = x]$. In fact, for any x,

$$\mathbb{P}[X = x] = \mathbb{P}[X \leq x] - \mathbb{P}[X < x] = 0 \; !$$

continuous random variables

- for continuous rv X with pdf $f(\cdot)$ and cdf $F(\cdot)$, we have

$$\mathbb{P}[a < X \leq b] = F(b) - F(a) = \int_a^b f(x)dx$$

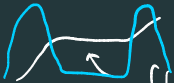
- now we can go from one function to the other as

$$F(x) = \int_{-\infty}^{x} f(z)\,dz$$

$$f(x) = \frac{dF(x)}{dx}$$

is defined as F is continuous

Eg - If $F(x)$ is 'flat'



$f(x)=0$

ORIE 4580/5580: Simulation Modeling and Analysis

ORIE 5581: Monte Carlo Simulation

unit 2: mean, variance, and tails

Sid Banerjee

School of ORIE, Cornell University

Why do we need this for simulation?

- Typical sim workflow: want to estimate some prob $\mathbb{P}[X \geq a]$
  - Design simulator Y s.t. $\mathbb{E}[Y] = \mathbb{P}[X \geq a]$
  - Generate independent copies $Y_1, Y_2, \ldots, Y_n$ for some chosen n

    replicates

    # of replications

expectations and independence

- Compute estimate $\hat{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$

  This 'works' because of Law of Large Numbers
- Note: $\hat{Y}$ is still random $\Rightarrow$ need 'confidence intervals'

  This 'works' because of Central Limit Thm

expected value (mean, average)

let X be a random variable, and $g(\cdot)$ be any real-valued function

- If X is a discrete rv with $\Omega = \mathbb{Z}$ and pmf $p(\cdot)$, then

$$\mathbb{E}[X] = \sum_{x} x\, p(x)$$

$$\mathbb{E}[g(X)] = \sum_{x} g(x)\, p(x)$$

$$\boxed{\begin{array}{l} \underline{\text{Looking Ahead}} \\ \mathbb{E}[\,\cdot\,] \text{ is LINEAR} \\ \text{ie } \mathbb{E}[a g(X) + h(X)] = a\,\mathbb{E}[g(X)] + \mathbb{E}[h(X)] \end{array}}$$

- if X is a continuous rv with $\Omega = \mathbb{R}$ and pdf $f(\cdot)$, then

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} z\, f(z)\, dz$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(z)\, f(z)\, dz$$

variance and standard deviation

Convention - uppercase $\Rightarrow$ r.v
lowercase $\Rightarrow$ number

$\mu_x = \mathbb{E}[X]$

$\sigma(X) = \sqrt{Var(X)}$

- definition: $Var(X) = \mathbb{E}\Big[\big(\overbrace{X - \underbrace{\mathbb{E}[X]}_{\text{is a number!}}}^{g(x)}\big)^2\Big]$
- (alternate formula for computing variance)

$$
\begin{aligned}
Var(X) &= \mathbb{E}\left[(X-\mu_x)^2\right] \\
&= \mathbb{E}\left[X^2 - 2\mu_x X + \mu_x^2\right] \\
&= \mathbb{E}[X^2] - 2\mu_x \underbrace{\mathbb{E}[X]}_{\mu_x} + \mu_x^2 \\
&= \mathbb{E}[X^2] - \mu_x^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2
\end{aligned}
$$

by the linearity of expectations!

Note - This means that $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$

independence

what do we mean by "random variables X and Y are independent"?
(denoted as $\underline{\underline{X \perp\!\!\!\perp Y}}$; similarly, $X \not\perp\!\!\!\perp Y$ for 'not independent')

intuitive definition: knowing X gives no information about Y

formal definition: $\mathbb{P}[X \leq x \text{ AND } Y \leq y] = \mathbb{P}[X \leq x]\,\mathbb{P}[Y \leq y] = F_X(x)\,F_Y(y)$

(assuming $X, Y \in \mathbb{R}$) FOR EVERY PAIR $(x,y) \in \mathbb{R}^2$

- one measure of independence between rv is their covariance

$$Cov(X,Y) = \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right] \qquad \mu_X = \mathbb{E}[X],\ \mu_Y = \mathbb{E}[Y] \qquad \text{(formal definition)}$$

$$= \mathbb{E}\left[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y\right] \qquad \text{(for computing)}$$

$$= \underbrace{\mathbb{E}[XY]}_{=\mathbb{E}[X]\mathbb{E}[Y] \text{ IF } X \perp\!\!\!\perp Y} - \mu_X \underbrace{\mathbb{E}[Y]}_{\mu_Y} - \mu_Y \underbrace{\mathbb{E}[X]}_{\mu_X} + \mu_X\mu_Y = \mathbb{E}[XY] - \mu_X\mu_Y$$

independence and covariance are confusing!!

how are independence and covariance related?

- X and Y are independent, then they are uncorrelated
  in notation: $X \perp\!\!\!\perp Y \quad \Rightarrow \quad Cov(X,Y) = 0$ Always true
- however, uncorrelated rvs can be dependent -> they can be anything...
  in notation: $Cov(X,Y) = 0 \quad \not\Rightarrow \quad X \perp\!\!\!\perp Y$
- $Cov(X,Y) = 0 \Rightarrow X \perp\!\!\!\perp Y$ only for multivariate Gaussian rv
  (this though is confusing; see this Wikipedia article)