* Expectations - mean, variance
* Law of large numbers
* Central limit thm - confidence intervals
- more general 'tail' bounds (Chebyshev's Ineq)

ORIE 4580/5580: Simulation Modeling and Analysis

ORIE 5581: Monte Carlo Simulation

Unit 3: expectation, independence, convergence

Sid Banerjee

School of ORIE, Cornell University

expectations and independence

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

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

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

$$\left(\text{Eg. } g(x) = (x - \mathbb{E}[X])^2 \right.$$
$$\left. \Rightarrow \mathbb{E}[g(X)] = \mathrm{Var}(X) \right)$$

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

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

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

variance and standard deviation

- **Definition**: $Var(X) = \mathbb{E}\Big[\underbrace{(X - \overset{\text{a number}}{\mathbb{E}[X]})^2}_{g(x)}\Big]$ 　 $\sigma(X) = \underbrace{\sqrt{Var(X)}}_{\text{Std-deviation}}$
- (More useful formula for computing variance)

$$
\begin{aligned}
Var(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\
&= \mathbb{E}\left[\left(X^2 - 2X\,\mathbb{E}[X] + \mathbb{E}[X]^2\right)\right] \\
&= \mathbb{E}[X^2] - 2\,\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\
&= \underbrace{\mathbb{E}[X^2] - \mathbb{E}[X]^2}_{\geq 0}
\end{aligned}
$$

Side-fact

$$\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$$

Why? because $g(x) \geq 0$

Universal property!!

independence

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

intuitive definition: knowing X gives no information about Y

formal definition: $\mathbb{P}[X \le x, Y \le y] = \underbrace{F(x)}_{\mathbb{P}[X \le x]} \cdot \underbrace{F(y)}_{\mathbb{P}[Y \le y]} \quad \forall x, y \in \mathbb{R}$

centering

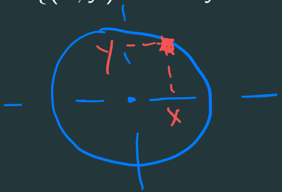
- One measure of ~~in~~dependence between rv is their covariance

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

$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \qquad \text{(for computing)}$$

clicker question

(X, Y) are uniformly distributed around unit circle $\{(x, y) : x^2 + y^2 = 1\}$



12

(a) $X \perp\!\!\!\perp Y$ and $Cov(X, Y) \neq 0$

22

(b) $X \perp\!\!\!\perp Y$ and $Cov(X, Y) = 0$

42

(c) $X \not\!\perp\!\!\!\perp Y$ and $Cov(X, Y) \neq 0$

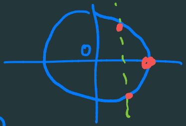
24

(d) $X \not\!\perp\!\!\!\perp Y$ and $Cov(X, Y) = 0$

clicker question: solution

(X, Y) are uniformly distributed around unit circle $\{(x, y) : x^2 + y^2 = 1\}$

- Independence - $X \not\perp\!\!\!\perp Y$



Suppose $\underline{X = 1}$ - what is Y?

- Covariance - $\mathrm{Cov}(X, Y) = 0$

- Note - $\mathbb{E}[X] = \mathbb{E}[Y] = 0$

$$\Rightarrow \mathrm{Cov}(X, Y) = \mathbb{E}[XY] = \int_{x=-1}^{1} \mathbb{E}[Yx] f(x)\, dx$$

$$= 0$$

independence and covariance

how are independence and covariance related?

- X and Y are independent, then they are uncorrelated

  in notation: $X \perp\!\!\!\perp Y \quad \Rightarrow \quad Cov(X, Y) = 0$ $(\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y])$

- however, uncorrelated rvs can be dependent

  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) !!

clicker question

TAs get lazy and distribute graded hws among 10 students uniformly at random; on average, how many students get their own hw?

(a) 0

(b) 0.1

(c) 1

(d) 2

(e) 3.1415...

clicker question

TAs get lazy and distribute graded hws among 10 students uniformly at random; on average, how many students get their own hw?

hint: indicator rvs + Linearity

Let $X_i = \mathbb{1}\left[\text{student i gets her hw}\right]$ (Indicator rv) $= \begin{cases} 1 & \text{if } i \text{ gets own hw} \\ 0 & \text{otherwise} \end{cases}$

$N =$ number of students who get their own hw $= \sum_{i=1}^{10} X_i$

Want $\boxed{E[N]}$

(a) 0

(b) 0.1

70%

(c) 1

(d) 2

(e) 3.1415...

clicker question: solution

TAs get lazy and distribute graded hws among 10 students uniformly at random; on average, how many students get their own hw?

Hint: Indicator random variables

Let $X_i = \mathbb{1}_{[\text{student i gets her hw}]}$ (Indicator rv)

$N =$ number of students who get their own hw $= \sum_{i=1}^{10} X_i$

$$\mathbb{E}[N] = \mathbb{E}\left[\sum_{i=1}^{10} X_i\right] = \sum_{i=1}^{10} \underbrace{\mathbb{E}[X_i]}$$

$$X_i = \begin{cases} 1 & \text{wp } 1/10 \\ 0 & \text{ow} \end{cases}$$

$$= \sum_{i=1}^{10} 1/10 = 1$$

Note - X_i are not independent! However, that does not matter for this because of linearity of expectation

linearity of expectation

for any rvs X and Y, and any constants $a, b \in \mathbb{R}$ - $\mathbb{E}[Z]$, $Z = aX + bY$

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

note 1: no assumptions! (in particular, does not need independence)

universal!! always true (for any r.v.)

linearity of expectation

for any rvs X and Y, and any constants $a, b \in \mathbb{R}$

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

$(*)$

note 1: no assumptions! (in particular, does not need independence)

note 2: does not hold for variance in general

- for general X, Y

  $$Var(aX + bY) = a^2 \, Var(X) + b^2 \, Var(Y) + 2ab \, Cov(X, Y)$$

- when X and Y are independent

  $$Var(aX + bY) = a^2 \, Var(X) + b^2 \, Var(Y)$$

  $(*)$

variance of linear combinations

Pf. $\mathrm{Var}(aX+bY) = \mathbb{E}\left[(aX+bY)^2\right] - \left(\mathbb{E}[aX+bY]\right)^2$

$$= \mathbb{E}\left[a^2X^2 + b^2Y^2 + 2abXY\right] - \left(a\,\mathbb{E}[X] + b\,\mathbb{E}[Y]\right)^2$$

$$\overset{\text{LOI}}{=} a^2\,\mathbb{E}[X^2] + b^2\,\mathbb{E}[Y^2] + 2ab\,\mathbb{E}[XY]$$

$$- a^2(\mathbb{E}[X])^2 - b^2(\mathbb{E}[Y])^2 - 2ab\,\mathbb{E}[X]\,\mathbb{E}[Y]$$

$$= \underbrace{a^2\,\mathrm{Var}(X)}_{a^2\left(\mathbb{E}[X^2] - (\mathbb{E}[X])^2\right)} + b^2\,\mathrm{Var}(Y) + \underbrace{2ab\,\mathrm{Cov}(X,Y)}_{2ab\left(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\right)}$$

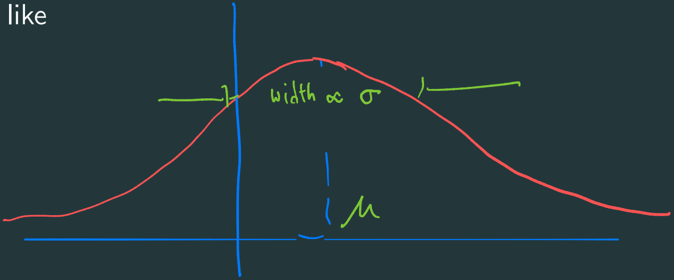
normal distribution

normal distribution

rv X is said to be normally distributed with mean μ and variance σ^2 (in notation, $X \sim \mathcal{N}(\mu, \sigma^2)$) if its pdf $f(\cdot)$ is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty.$$

the pdf looks like



properties of the normal distribution

1. the pdf is symmetric around the mean μ: if $X \sim \mathcal{N}(\mu, \sigma^2)$

   $\mathbb{P}[X \leq \mu - a] =$

2. (Linear transformation) If $X \sim \mathcal{N}(\mu, \sigma^2)$, then

   $$aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

   $$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

   'standardizing'

3. If $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$, and $X \perp\!\!\!\perp Y$, then

   $$X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

cdf of normal distribution

if $X \sim \mathcal{N}(\mu, \sigma^2)$, then its cdf is given by

$\mathbb{P}[X \leq x] =$ $$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2} dy$$

- knowing cdf for $\mathcal{N}(0,1)$ is enough to find cdf for any normally distributed rv
- for $X \sim \mathcal{N}(0,1)$, its cdf is denoted $\Phi(x)$, and available in most computing packages.

  it is also closely related to the error function $\underline{erf(x)} = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} dt$;

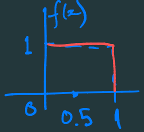
  in particular,

$$\Phi(x) = \frac{1}{2}(1 + erf(\frac{x}{\sqrt{2}}))$$

sums of independent rv

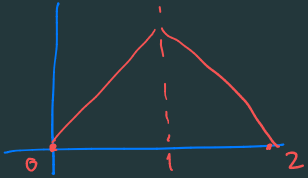
sums and averages of independent rv

- $X_1, X_2, \ldots$ are independent rv that are uniformly distributed over the interval $[0,1]$; $\mathbb{E}[X_1] = 1/2$, $Var(X_1) = 1/12$.
- the pdf of X_1 looks like



- what about the pdf of $X_1 + X_2$ and $(X_1 + X_2)/2$?

$X_1 + X_2$



sums and averages of independent rv

X_i iid, $\mathbb{E}[X_i] = \mu$

$Var(X_i) = \sigma^2$

$$S_n = X_1 + \ldots + X_n \qquad \bar{X}_n = \frac{1}{n}[X_1 + \ldots + X_n]$$

- $\mathbb{E}[S_n] =$ $n\mu$

  $\mathbb{E}[\bar{X}_n] =$ μ

  $Var(S_n) =$ $n\sigma^2$

  $Var(\bar{X}_n) =$ σ^2/n

Recipe to get $N(0,1)$
- Subtract the mean $\left(\sum_i (X_i - \mu)\right)$
- Divide by $\sqrt{n}\,\sigma$ $\left(\sum_{i=1}^{n} \frac{(X_i-\mu)}{\sqrt{n}\,\sigma}\right)$

$$\left(\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{n}{n^2}\sigma^2 = \frac{\sigma^2}{n}\right)$$

- (roughly) sum of n i.i.d. random variables is $\sqrt{n}$ times as variable as any one of the random variables
- average of n i.i.d. random variables is $1/\sqrt{n}$ times as variable as any one of the random variables

law of large numbers

let $X_1, X_2, \ldots$ be a sequence of independent rvs with $\mathbb{E}[X_i] = \mu$ for all i

then, "almost" always

$$\underbrace{\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i}_{\text{rv}} \longrightarrow \underbrace{\mu}_{\text{number}}, \quad \text{as } n \to \infty$$

note: for any finite n, $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ is still a random variable

central limit theorem

• $Z_1, Z_2 \sim N(0,1)$, $\boxed{\frac{Z_1+Z_2}{\sqrt{2}} \sim N(0,1)}$

let $X_1, X_2, \ldots$ be a sequence of independent rvs with

$$\mathbb{E}[X_i] = \mu, \; Var(X_i) = \sigma^2 < \infty \text{ for all } i$$

then, $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$, $\mathbb{E}[Z_n] = 0$, $Var(Z_n) = 1$

$$\underbrace{\overbrace{\sqrt{n}(\bar{X}_n - \mu)}^{}}_{\text{r.v.}} \xrightarrow{D} \underbrace{\sigma\mathcal{N}(0,1)}_{\text{r.v.}} = \underbrace{\mathcal{N}(0,\sigma^2)}_{\text{distrib}}, \quad \text{as } n \to \infty$$

r.v. ⟷ r.v. ∼ distrib

this suggests the following approximations for large n,

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \overset{D}{\approx} N\left(\mu, \frac{\sigma^2}{n}\right)$$

("roughly" speaking)

not formal as these do not 'converge' to any distribution. See notebook for visualization

$$S_n = \sum_{i=1}^{n} X_i \overset{D}{\approx} N(n\mu, n\sigma^2)$$