Review of Probability

probability space $(\Omega, \mathcal{F}, P)$

Ω : set of "outcomes". finite or infinite.

$\mathcal{F} \subseteq 2^{\Omega}$: set of "events". $\mathcal{F} \neq \phi$.

$\mathcal{F}$ is a <u>σ-algebra</u>:

① : $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.

② : A_i : countable sequence of sets

$\Rightarrow \bigcup A_i \in \mathcal{F}$.

$P : \mathcal{F} \to [0,1]$:

① $P(\phi) = 0$.

② A_i : countable sequence of <u>disjoint</u> sets

$$P(\bigcup A_i) = \sum_i P(A_i).$$

Example. Discrete space. Ω is countable.

$\mathcal{F} = 2^{\Omega}$. $\hat{p} : \Omega \to [0,1] : \sum_{\omega \in \Omega} \hat{p}(\omega) = 1$.

$\forall A \in \mathcal{F}$: $P(A) \triangleq \sum_{\omega \in A} \hat{p}(\omega)$.

Question. How to define a probability on $\mathbb{R}$?

Or what do we mean by drawing a uniform real in $(0,1)$.

$\Omega = (0,1)$

$\mathcal{F} =$ "the Borel Sets"

" Countable union / complement of all open intervals "

P: $\forall$ interval $I = (a,b)$, $P((a,b)) = b-a$

Lebesgue measure.

Random Variable

Given $(\Omega, \mathcal{F}, \mathbb{P})$.

real-valued random variable:

$$X : \Omega \to \mathbb{R}$$

Countable Ω:

$\forall a \in \mathrm{Ran}(X)$:

$$\mu(a) = \Pr[X = a] \triangleq \mathbb{P}[\{w : X(w) = a\}]$$

$$= \mathbb{P}[X^{-1}(a)].$$

Example Binomial.

$$X \sim \mathrm{Binom}(n, p).$$

$$(\Rightarrow) \quad \Pr[X = k] = \binom{n}{k} \cdot p^k (1-p)^k.$$

$\Omega = \{0, 1\}^n$. ...

Example. Geometric. $X \sim \text{Geometric}(p)$.

$$\Pr[X = k] = (1-p)^{k-1} p.$$

$$\Omega = \{0,1\}^*.$$

Uncountable Ω.

density function f of X:

$$\Pr(a \le X \le b) = \int_a^b f(x)\,dx.$$

distribution function F:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt. = \Pr(-\infty < X \le x).$$

Example. Uniform on (a,b).

$$f(x) = \begin{cases} 1/b-a, & a < x < b. \\ 0, & \text{o.w.} \end{cases}$$

Example. Exponential distribution with $\lambda > 0$.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0. \\ 0, & \text{otherwise.} \end{cases}$$

Example. Gaussian / Normal.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Expectation of X.

discrete $\quad E[X] = \sum_{a} a \cdot P_r[X=a]$

continuous $\quad E[X] = \int_{-\infty}^{\infty} t \cdot f(t)\, dt.$

uniform: $\quad \int_a^b t \cdot \frac{1}{b-a}\, dt = \frac{1}{b-a} \cdot \frac{t^2}{2} \Big|_a^b = \frac{b+a}{2}.$

exponential: $\quad \int_0^{\infty} t \cdot \lambda e^{-\lambda t}\, dt = -\int t\, d e^{-\lambda t}$

$$= -t e^{-\lambda t} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda t}\, dt = -\frac{1}{\lambda} e^{-\lambda t} \Big|_0^{\infty}$$

$$= \frac{1}{\lambda}.$$

Variance of X:

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}X^2 - (\mathbb{E}X)^2.$$

Independence.

$X \perp Y$ if $\forall A, B$. $\Pr[X \in A \wedge Y \in B] =$

$\Pr[X \in A] \cdot \Pr[Y \in B]$.

Linearity of Expectations

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i].$$

$X_1, \dots, X_n$: "Mutually independent"

$$\mathbb{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathbb{E}[X_i].$$

$$\mathrm{Var}\left[\sum x_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i].$$

Conditional Probability in $(\Omega, \mathcal{F}, P)$

Two events $A, B \in \mathcal{F}$.

$$P[A \mid B] \doteq P[A \cap B] / P[B].$$

Measurable

A probability space $(\Omega, \mathcal{F}, P)$.

$X : \Omega \to \mathbb{R}$.

Discrete. $\forall a \in \mathrm{Ran}(X)$. $X^{-1}(a) \in \mathcal{F}$.

X is $\mathcal{F}$-measurable.

$\sigma(X)$: the minimal σ-algebra $\mathcal{F}$ such that

X is $\mathcal{F}$-measurable.

$\forall X, Y : \Omega \to \mathbb{R}$.

$$\forall a : \quad \mathbb{E}[X \mid Y = a] = \sum_b b \cdot \Pr[X = b \mid Y = a]$$

$$f(Y) = \mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)]$$

$$\forall \omega \in \Omega: \quad f(Y)(\omega) = \mathbb{E}[X \mid Y = Y(\omega)]$$

<u>Prop</u>. ①. $\mathbb{E}[X \mid Y]$ is $\sigma(Y)$ - measurable.

② $\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[f(Y)] = \mathbb{E}[X]$.

Pf. ① is trivial.

②. $\mathbb{E}[f(Y)] = \sum_a \mathbb{E}[X \mid Y = a] \cdot \Pr[Y = a]$

$$= \sum_a \sum_b b \cdot \Pr[X = b \mid Y = a] \cdot \Pr[Y = a]$$

$$= \sum_b b \cdot \sum_a \Pr[X = b \mid Y = a] \cdot \Pr[Y = a]$$

$$= \sum_b b \cdot \sum_a \Pr[X = b \wedge Y = a] = \sum_b b \cdot \Pr[X = b]$$

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

X : height of a person.

Y : gender of a person.

$(\Omega, \mathcal{F}, P)$. Ω - Chinese people.

$\mathbb{E}[X]$ - average height.

$f(Y) = \mathbb{E}[X \mid Y] : \Omega \to \mathbb{R}$.

$f(Y)(w)$: the average height of people with gender same as w.

$$\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X].$$