Review of Probability

Experiments. The sample space. "S"

$A \subseteq S$ $\quad$ $A \rightarrow$ event

$\Phi \subseteq S$ $\quad$ $S \subseteq S$

Probability:

1. $0 \leq P(A) \leq 1$
2. $P(\Phi) = 0$, $P(S) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, A_i are pairwise disjoint

Examples: 4.1.2 Flip a coin three times.

Sample space $= \{$ ~~TTT, HHH, THH, TTH~~ $\}$

$= \{$TTT, TTH, THT, THH,

HTT, HTH, HHT, HHH }

<u>Event</u>: {Two Tails, One Head} → 3 $P(A) = \frac{3}{8}$

{At least one tail} → $P(*) = \frac{7}{8}$

<u>Example</u> 4.1.3 $\{t : \underline{0} \leq t \leq 24\}$

<u>Example</u>: Height of student at Temple.

Height Randomness ⟺ Random Sampling

Conditional Probability. , A and B

$P(\underline{A \mid B})$ <u>Example</u>: Toss a die.

$\underline{A} = \{\text{I win}\} = \{\text{Value} \leq 3\}$ $P(A) = \frac{1}{2}$

$B = \{\text{odd number}\}$.

$P(\bar{A} \mid B)$

$P(A \mid B) = \frac{P(A \cap B)}{\quad}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A|B) = \frac{P(1,3)}{P(1,3,5)} = \frac{2/6}{3/6} = \frac{2}{3}$$

Disease:	Present	Absent
Test	T+	T−

Two-way Contingency Table

	D	N	
T+	950	10	960
T−	50	990	1040
	1000	1000	2000

$$P(T_+ | D) = \frac{950}{1000} = 0.95$$

$$P(T_+ | D) = \frac{P(T_+ \cap D)}{P(D)} = \frac{950/2000}{1000/2000} = 0.95$$

Multiplicative law:

$$P(A \cap B) = P(B) \cdot P(A|B)$$

$$\neq P(B) \cdot P(A)$$

Independent $\quad P(A|B) = P(A)$

$\underline{D} \qquad \underline{T_+}$

$$P(T_+|D) = 0.95$$

$$P(T_+) = \frac{960}{2000} = 0.48 \neq 0.95 \Rightarrow D \text{ and } T_+ \text{ are not independent.}$$

Gender vs. Right-handed / Left handed

	RH	LH	
Male	44	11	55
Female	36	9	45
	80	20	100

| 80 | 20 | 100

$P(\cancel{MALE \mid LH})$ $P(RH \mid Male) = \frac{44}{55} = \frac{4}{5} = 0.8$

$$P(RH) = \frac{80}{100} = 0.8$$

Independent

	RH	LH	
M	45	10	55
F	36	9	45
	81	19	100

$$P(RH \mid M) = \frac{45}{55} = \frac{9}{11}$$

$$P(RH) = \frac{81}{100} = 0.81 \neq \frac{9}{11} \Rightarrow$$

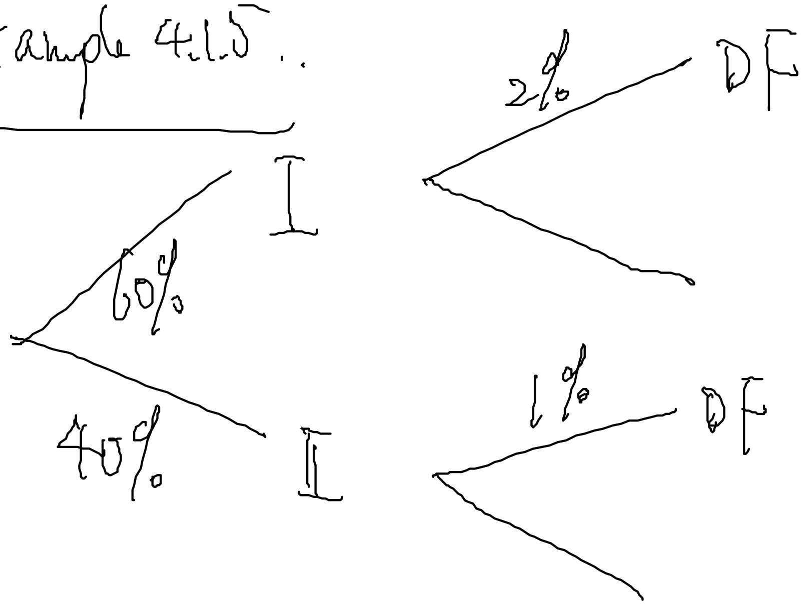
Bayes Theorem:

Let $B_1 \dots B_n$ be a partitions of S. for any event

A_i

$$P(B_i|A) = \frac{P(A|B_i)\,P(B_i)}{\sum_j P(A|B_j)\,P(B_j)}$$

Example 4.1.5 ..



$$P(I|DF) = \frac{P(DF|I)\,P(I)}{P(DF|I)\,P(I) + P(DF|II)\,P(II)}$$

$$= \frac{2\% \times 60\%}{2\% \times 60\% + 1\% \cdot 40\%} = \frac{120}{120+40} = \frac{3}{4} = 0.75$$

Example: 1,000 athletes, 10 are illegally drug user

Positive 95% if drug-User,

Negative 95% if non DU

0.05, 99

$P(\text{Not DR} \mid \text{Positive})$ 55%

$$= \frac{P(\text{Positive} \mid \text{Not DR}) \cdot P(\text{Not DR})}{P(\text{Positive} \mid \text{Not DR})\,P(\text{Not}) + P(\text{Positive} \mid \text{DR})\,P(\text{DR})}$$

$$= \frac{0.05 \times 0.99}{0.05 \times 0.99 + 0.95 \times 0.01} = 83.9\%.$$

4.1.3 $H \to$ Hypothesis. Model
Theory

$D \longmapsto$ Data.

$$P(D \mid H)$$

Statistical Inference : Based on the data, infer about the underlying model.

$$P(H \mid D)$$

Random Variables R.V.

$X : S \mapsto \mathbb{R}$

$\mathbb{R} \supseteq \{1, -1\}$

$X_i : S \mapsto \{1, -1\}$

$1 \to 1$
$2 \to 1$
$3 \to 1$
$4 \to -1$
$5 \to -1$
$6 \to -1$

Discrete / Continuous.

possible value is countable

uncountable forms an interval

Discrete r.v.

Bernoulli (p): ~~toss~~ toss a coin, p - head

Let X be the number of head

$X: \{0, 1\}$

$X \sim Ber(p)$

$P(X=0) = 1-p, \quad P(X=1) = p.$

Binomial Distribution

Toss the same coin n times,

X = total number of heads

$k = 0, 1, 2, \cdots n$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$X \sim Bin(n, p)$

Poisson Distribution

Let X be the number of independent events per unit time, $X \sim Pois(\lambda)$

X, 0, 1, 2, ... ∞

$$P(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda) \quad k = 0, 1, 2, \cdots$$

~~Pdf~~ Probability Mass Function (pmf)

Example: A life insurance salesmen sells on average 3 policies in a week.

① What is the prob. that he/she will sell some policies in the next week?

② Assuming he/she works five days a week. What is the prob. that he/she will sell one policy on next Monday?

Let: X be the number of policies he/she will sell in the next week.

$$X \sim \text{Poisson}(3)$$

$$P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + \cdots$$

$$= 1 - P(X=0) = 1 - \frac{3^0}{0!}\exp(-3) = 1 - e^{-3}$$

Let Y be the number of policies he/she will sell in one day

$$Y \sim \text{Poisson}\left(\frac{3}{5}\right)$$

$$P(Y=1) = \frac{\lambda^1}{1!}e^{-\lambda} = \frac{3}{5}e^{-\frac{3}{5}}$$

$$P(X = x)$$

Continuous R.V.

$$P(X = x) = 0$$

$P(\underline{X = 1.72}) = 0$ $\qquad$ $P(X \leq 1.72)$

Cumulative Distribution ~~function~~ function. (CDF)

$\underline{F_X(x)} = P(\underline{X} \leq x)$

probability density function (pdf)

$$f_{\underline{X}}(x) = \left(F_X(x)\right)'$$

① $f_X(x) \geq 0$

② $\int_{-\infty}^{\infty} f(x)\,dx = 1$ ✓

<u>Uniform</u> Distribution:

$X \sim U(a, b)$ if $f_X(x) = \frac{1}{b-a}$

If $a=0$, $b=1$.

$X \sim U(0,1)$ if $f_X(x) = 1$.

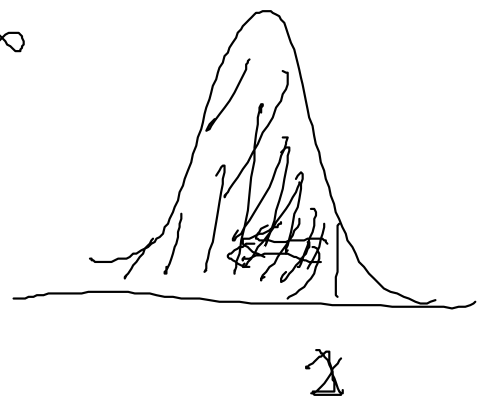
Normal Distribution

$$X \sim N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

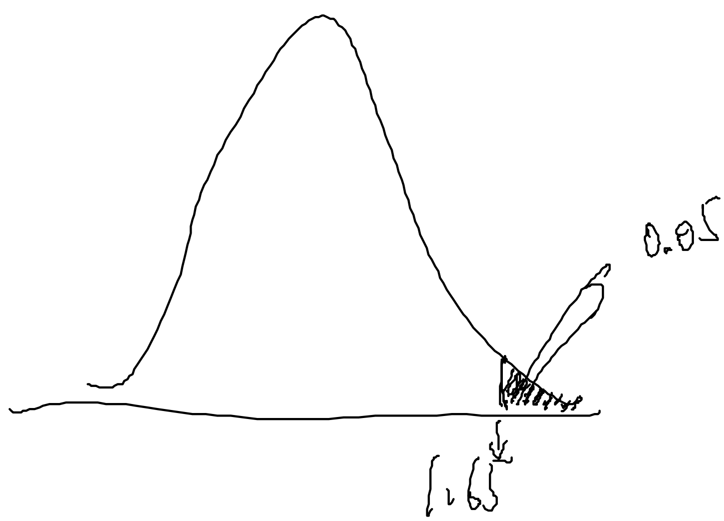
$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx = \Phi(x)$$

$\Phi(x)$



Critical values

Quantile



Exponential $X \sim Exp(\lambda)$ if

$$f_X(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \qquad x > 0$$

$$F_X(x) = \int_0^x f_X(x) = 1 - e^{-\frac{x}{\lambda}}$$

dexp rexp. qexp pexp

In R: dexp ... $\lambda^* = \frac{1}{\lambda}$

$$f_X(x) = \lambda e^{-\lambda x}.$$

Gamma Distribution. $\to$ shape $\to$ scale rate $= \frac{1}{\lambda}$.

$X \sim \Gamma(\alpha, \lambda)$ if

$$f(x) = \frac{1}{\lambda^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\lambda}} \qquad x > 0$$

Inverse Gamma Distributn $X \sim \Gamma(\alpha, \lambda)$

$$Y = \frac{1}{X} \sim \text{INV Gamma}(\quad)$$

prior for the variance