IES302 Review

Wednesday, January 18, 2012
1:29 PM

Review

- IES 302
  - Part I → Probability
  - Part II → Statistics

Section 1

- Long-term relative frequency interpretation of probability:

  If an event A has probability P(A),

  then suppose you can repeat the experiments many times,

  the proportion of times that A occurs will be approximately P(A).

  relative frequency

Section 3

- Classical Probability:

  1. Assumptions
     - Finite Ω
     - Equipossibility

  2. $P(A) = \dfrac{|A|}{|\Omega|}$ → Finding the cardinality (size) of sets will be important.

Section 4

- Counting Techniques
  - principles/rules
    - addition (cases)
    - multiplication (steps)
    - subtraction

("at least")

division

Four Types of Sampling Problems

Ordered / unordered sampling — with replacement / w/o replacement

OSwR : n^r

OSwoR : $(n)_r$ ($\cancel{P^n_r = {}^nP_r}$)

USwoR : $\binom{n}{r}$

USwR : → study/read notes if u r interested.

Section 5

General (event-based) Probability Theory

- Kolmogorov's Axioms

P1: $P(A) \geq 0$

P2: $P(\Omega) = 1$

P3: $P(A \cup B) \overset{A \perp B}{=} P(A) + P(B)$

$P(A_1 \cup A_2 \cup \cdots \cup A_n)$

$\overset{\perp}{=} P(A_1) + P(A_2) + \cdots + P(A_n)$

The sets involved must be disjoint

$P(A_1 \cup A_2 \cup \cdots) \overset{\perp}{=} P(A_1) + P(A_2) + P(A_3) + \ldots$

Properties (derived from Axioms)

$P(\emptyset) = 0$

$0 \leq P(A) \leq 1 \qquad (P(A) \in [0,1])$

$P(\{a_1, a_2, \ldots, a_n\}) = P(\{a_1\}) + P(\{a_2\}) + \cdots + P(\{a_n\})$

$$P(\{a_1, a_2, \dots, a_n\}) = P(\{a_1\}) + P(\{a_2\}) + \dots + P(\{a_n\})$$

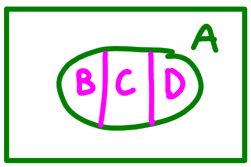
$$P(\{a_1, a_2, \dots\}) = P(\{a_1\}) + P(\{a_2\}) + \dots$$

"Venn diagram"-based formulas

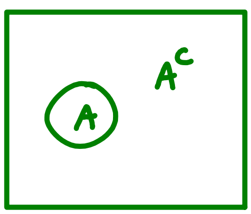
useful when you have a couple of sets

Think of the Ω as having area 1.

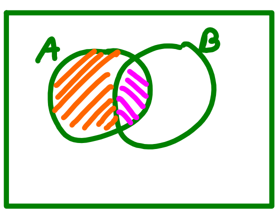
Finding probability $\Rightarrow$ Finding area.



$$P(A) = P(B) + P(C) + P(D)$$



$$P(A) + P(A^c) = 1$$



$$P(A) = P(A \setminus B) + P(A \cap B)$$

$$P(A \setminus B) = P(A) - P(A \cap B)$$

$$P(A \cup B) = P(B) + P(A \setminus B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

In general

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) \leq P(A_1) + P(A_2) + \dots + P(A_n)$$

$$P(A_1 \cup A_2 \cup \dots) \leq P(A_1) + P(A_2) + \dots$$

Each inequality above becomes equality when the sets involved in the union are disjoint.

Section 6.1: Conditional Probability

Def. $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

"Probability of A given B"

$$P(A^c|B) = 1 - P(A|B)$$

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

$$P(A|C \cap D) = P(A|C, D) = \frac{P(A \cap C \cap D)}{P(C \cap D)}$$

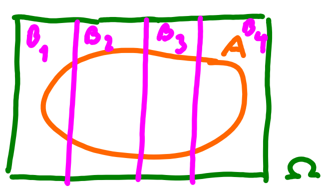
comma means "AND"

$$P(A \cap B \cap C) = P(C|\underbrace{A, B}_{A \cap B})\,P(B|A)\,P(A)$$

$$P(A \cap B|C) = P(A|B, C)\,P(B|C)$$

Total Probability Theorem:

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)\,P(B_i)$$



Bayes' Theorem

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

use total probability theorem to calculate $P(A)$

Section 6.2 Independence

Def. $A \perp\!\!\!\perp B$ iff $P(A \cap B) = P(A)P(B)$

Three more equivalent statements

$A^c \perp\!\!\!\perp B$, $A \perp\!\!\!\perp B^c$, $A^c \perp\!\!\!\perp B^c$

$\Updownarrow$

$P(A^c \cap B) = P(A^c)P(B)$

$\Updownarrow$

$P(A^c \cap B^c) = P(A^c)P(B^c)$

When we know that $P(A) \neq 0$ and $P(B) \neq 0$, then $A \perp\!\!\!\perp B$ is also equivalent to

$P(A|B) = P(A)$

$P(B|A) = P(B)$

Def. A, B, C are independent iff all four of the following conditions hold

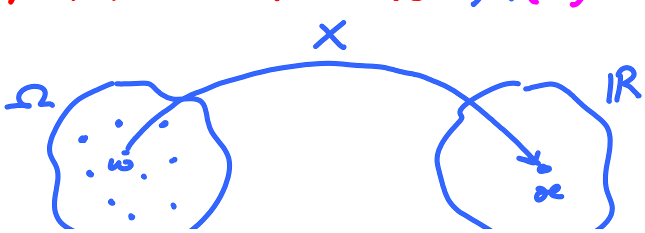
$$\left.\begin{array}{l} 1)\ A \perp\!\!\!\perp B \\ 2)\ A \perp\!\!\!\perp C \\ 3)\ B \perp\!\!\!\perp C \end{array}\right\} \Leftrightarrow \text{pairwise independence}$$

4) $P(A \cap B \cap C) = P(A)P(B)P(C)$

Bernoulli Trial : two outcomes $\left\langle \begin{array}{l}\text{success} \\ \text{failure}\end{array}\right.$

n Bernoulli trials (implicitly assume that the trials are independent)

Section 7 Random Variable $X(\omega) = x$



S_X = support of RV X
= any set such that $P[X \in S_X] = 1$

Three types of RVs
1) discrete
2) continuous
3) mixed

Section 8 Discrete RV

finite
countably infinite

X is discrete iff X has a countable support

iff you can find numbers
$x_1, x_2, \ldots, x_n$
or a sequence of numbers
$x_1, x_2, \ldots$
such that $\sum_k P[X = x_k] = 1$

Probability mass function (pmf)

$$p_X(x) = P[X = x]$$

For example,

$$p_X(2) = P[X = 2]$$

$$p_X(\sqrt{2}) = P[X = \sqrt{2}]$$

- To find

$$P[\text{some condition(s) on } X]$$

$P[X \geq 3]$, $P[X^2 = 4]$

from the pmf $p_X(x)$ of X :

from the pmf $p_X(x)$ of X :

step 0: Find the support of X

step 1 : Look only at x inside the support
find all x that satisfies the condition(s)

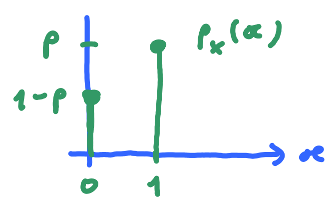
step 2 : Use x found in step 1,
evaluate the pmf at these points.

step 3 : Add the pmf values from step 2.

- Any pmf $p(\cdot)$ satisfies
  1) " ≥ 0 "
  2) " $\sum = 1$ "

- Ex. Bernoulli RV

$$p_X(x) = \begin{cases} 1-p, & x = 0, \\ p, & x = 1, \\ 0, & \text{otherwise}. \end{cases}$$



CDF : cumulative distribution function

$$F_X(x) = P[X \leq x]$$

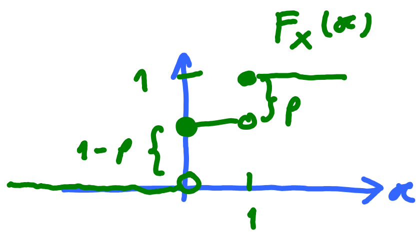
add $p_X(\cdot)$ upto (including) the point x.

For discrete RV X,

$F_X(x)$ is a right-continuous staircase function of x with jumps at x where pmf $p_X(x) > 0$.

this is the size of the corresponding jump.

Ex. CDF for Bernoulli RV X above:



Using the location/size of the jump(s), we can "read" pmf from cdf.

- General properties of CDF

  1) nondecreasing

  2) right-continuous

  3) $\lim_{x \to -\infty} = 0$, $\lim_{x \to \infty} = 1$

Families of discrete RVs.

Uniform on $\{1, \dots, n\}$ $\quad p_X(x) = \begin{cases} \frac{1}{n}, & x = 1, 2, \dots, n \\ 0, & \text{otherwise.} \end{cases}$

Bernoulli

Binomial $\quad p_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, \dots, n, \\ 0, & \text{otherwise.} \end{cases}$

Bernoulli trials

Each trial is Bernoulli RV with success probability $= p$.

$\frac{1}{2}$ for fair coin

success → 1

failure → 0

⟼ # trials until the next success ~ geometric

$\mathcal{G}_0$

Each trial is Bernoulli RV with success probability $= p$.

$\frac{1}{2}$ for fair coin

success $\to 1$

failure $\to 0$

g_0

trials until the next success ~ geometric

g_1

100 110 100 11 000101 000 1

successes in n trials ~ binomial (n, p)

as $n \to \infty$, $p \to 0$
(n large, p small)

Poisson(np)

Expectation (Expected value) "average"

$$\mathbb{E}X = \sum_x x\, p_X(x)$$

$$\mathbb{E}[g(X)] = \sum_x g(x)\, p_X(x) \qquad \mathbb{E}[X^4] = \sum_x x^4 p_X(x)$$

$$\mathbb{E}[aX + b] = a\,\mathbb{E}X + b \qquad \mathbb{E}[2X-1] = 2\,\mathbb{E}X - 1$$

Caution: In general, $\mathbb{E}[g(X)] \neq g(\mathbb{E}X)$

$$\mathbb{E}[X^4] \neq (\mathbb{E}X)^4$$

Extension:

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

$$\mathbb{E}[X - 5Y + 3] = \mathbb{E}X - 5\,\mathbb{E}Y + 3$$

$$\mathbb{E}[a_1 X_1 + a_2 X_2 + \cdots + a_n X_n + c]$$

$$\mathbb{E}[a_1 X_1 + a_2 X_2 + \cdots + a_n X_n + c]$$
$$= a_1 \mathbb{E}X_1 + a_2 \mathbb{E}X_2 + \cdots + a_n \mathbb{E}X_n + c$$

$$\mathbb{E}[a g_1(X) + b g_2(X)] = a \mathbb{E}[g_1(X)] + b \mathbb{E}[g_2(X)]$$

$$\mathbb{E}[5X^2 - 3X^3] = 5\mathbb{E}[X^2] - 3\mathbb{E}[X^3]$$

$$\mathbb{E}[a_1 g_1(X) + a_2 g_2(X) + \cdots + a_n g_n(X)]$$
$$= a_1 \mathbb{E}[g_1(X)] + a_2 \mathbb{E}[g_2(X)] + \cdots + a_n \mathbb{E}[g_n(X)]$$

$$\mathbb{E}[5X^2 - 3X^3 + 2\sqrt{X} - 7\sin X]$$
$$= 5\mathbb{E}[X^2] - 3\mathbb{E}[X^3] + 2\mathbb{E}[\sqrt{X}] - 7\mathbb{E}[\sin X]$$