Probability and statistics

August - 24

lecture #4

R: Random expt.

Ω: Sample space, $\mathcal{F}$: σ-algebra, P: probability measure.

In case, additional information is available about R, ($B \in \mathcal{F}$ occurrs)

conditional probability.

$$\underbrace{P(A \mid B)} = \frac{P(A \cap B)}{P(B)} \qquad P(B) > 0$$

$$P(A) = P(A \mid \Omega) = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1} = P(A)$$

<u>disjoint cover</u>: $\{B_1, \ldots, B_n\}$ disjoint cover of Ω

$$P(A) = P(A \mid B_1) P(B_1) + \cdots + P(A \mid B_n) P(B_n)$$

<u>Baye's rule</u>: $P(A \cap B) = P(A \mid B)\, P(B) = P(B \mid A)\, P(A)$

Independence of events.

Product rule: $P(A \cap B) = P(A|B)\,P(B)$ —(1)

$= P(B|A)\,P(A)$ —(2)

Definition: $A, B \in \mathcal{F}$ are said to be independent (statistically independence) if

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

From (1), $P(A|B) = P(A)$

From (2), $P(B|A) = P(B)$

Ex:

R : Random exp^{t}

$\Omega = \{1, 2, 3, 4\}$

Assumption: All outcomes are equally likely.

$A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 4\}$

$P(A) = \frac{2}{4} = \frac{1}{2}$

$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{2}$

$P(B) = \frac{1}{2}$

$P(B \mid A) = \frac{1}{2}$

$P(A \cap B) = \frac{1}{4} = P(A)\,P(B)$

A & B are independent events. $\blacksquare$

Pairwise independence of events

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

In the previous example A, B are pairwise independent and so are B, C and A, C.

$$P(A\cap B\cap C) = \frac{1}{4} \neq P(A)\cdot P(B)\cdot P(C) = \frac{1}{8}$$

Definition: Let A, B, C be three events from $(\Omega, \mathcal{F}, P)$. We call A, B, C mutually independent if

1) A, B, C are pairwise independent.
$P(A\cap B) = P(A)P(B)$, $P(B\cap C) = P(B)P(C)$, $P(A\cap C) = P(A)P(C)$

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

We can generalize the definition of mutual independence to the events $A_1, A_2, \dots, A_n$ in an inductive way.

- Any subcollection of $A_1, \dots, A_n$ containing atleast 2 events and atmost $n-1$ events is mutually independent.
- $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2) \dots P(A_n)$

Application of independence in setting up probability space in case of repeated random expt. with only 2 possible outcomes.

<u>Note:</u> A random expt. with only 2 outcomes is called as Bernoulli trial.

$\Omega = \{0, 1\}$

$\mathcal{F} = 2^{\Omega} = \{\{0\}, \{1\}, \phi, \{0, 1\}\}$

Assignment of probability : $P\{1\} = p \leftarrow$ probability of success.

$\Rightarrow P\{0\} = 1 - p$

<u>Aim</u>: Set up the probability space when the random expt. R/ Bernoulli trial is repeated 'n' number of times with the assumption that the trials /repetitions are <u>independent</u> of each other.

For the compound expt / n-reptition of Bernoulli trial.

$$\Omega_c = \{ \underbrace{y_1 y_2 \cdots y_n}_{} \mid y_1, y_2, \ldots, y_n \in \{0, 1\} \}$$

$0100 \cdots 011$, $11 \cdots 1$, $00 \cdots 0$

$\# \Omega_c = 2^n$ (finite set)

$\mathcal{F}_c$ = power set $\Omega_c = 2^{\Omega_c}$

$\{1010 \cdots 01\} \in \mathcal{F}_c$

$\{00 \cdots 0\} \in \mathcal{F}_c$

$\{111 \cdots 1\} \in \mathcal{F}_c$

$\{y_1 y_2 \cdots y_n\} \in \mathcal{F}_c$

singleton events

<u>Probability assignment:</u>

It is enough to assign probability to singleton events $\{y_1 y_2 \cdots y_n\}$.

Consider an element $\omega = y_1 y_2 \cdots y_n \in \Omega_c$ such that there are exactly k number of 1's at specified location.

Let us assume that all the 1's are in the first k position.

$y_1 = 1, y_2 = 1, \ldots, y_k = 1; \quad y_{k+1} = 0 = y_{k+2} = \cdots = y_n$

$$\omega = (\underbrace{1\,1 \cdots 1}_{k}\ \underbrace{0\,0 \cdots 0}_{n-k})$$

(There are exactly k number of successes in n independent Bernoulli trials at specified locations)

A_i = Success at i^{th} trial $\quad i = 1, 2, \ldots, n.$

$$p(\omega) = P(\underbrace{1\,1 \cdots 1}_{k}\ \underbrace{0\,0 \cdots 0}_{n-k}) = P(A_1 \cap A_2 \cap \cdots \cap A_k \cap A_{n+1}^c \cap \cdots \cap A_n^c)$$

$$= P(A_1) P(A_2) \cdots P(A_k) P(A_{k+1}^c) \cap \cdots \cap P(A_n^c)$$

$$= p^k (1-p)^{n-k}$$

Assumption of mutually independent Bernoulli trials

Consider event A_k = All outcomes with exactly k number of successes.

$$P(A_k) = \binom{n}{k} \underbrace{p^k (1-p)^{n-k}} \qquad k = 0, 1, 2, \cdots, n$$

$$A_k = \underset{\binom{n}{k}}{\bigcup} \{y_1 \cdots y_n\}$$

<u>Example:</u> A box contains $n = 10$ balls

6 - red (60%)

4 - blue (40%)

Any ball is equally likely to be picked

Case (1) R: Pick a ball and replace it in the urn and a pick a ball again. } sampling with replacement.

A_1 = red ball is picked in the first trial.

A_2 = blue ball is picked in the second trial.

$P(A_2 | A_1) = \frac{4}{10}$; $\boxed{P(A_2) = \frac{4}{10}}$

''

$P(A_2) = P(A_2|A_1)\,P(A_1) + P(A_2|A_1^c)\,P(A_1^c)$

case(2): R: Pick a ball; do not replace in the box and pick a ball again. } sampling without replacement.

A_1 = red ball in the first trial.

A_2 = blue ball in the second trial.

$P(A_2 \mid A_1) = \frac{4}{9}$; $P(A_2) = P(A_2 \mid A_1)P(A_1) + P(A_2 \mid A_1^c)P(A_1^c)$

$$= \frac{4}{9} \cdot \frac{6}{10} + \frac{3}{9} \cdot \frac{4}{10}$$

$$P(A_2) = \frac{4}{10}$$

$$P(A_2 \mid A_1) \neq P(A_2)$$

simple exercise: Sampling without replacement.

$n = 10$ (0.044) $P(A_2 \mid A_1) = \frac{4}{9}$; $P(A_2) = \frac{4}{10}$

$n = 100$ (0.004) $P(A_2 \mid A_1) = \frac{40}{99}$; $P(A_2) = \frac{40}{100}$

$n = 1000$ (0.0004) $P(A_2 \mid A_1) = \frac{400}{999}$; $P(A_2) = \frac{400}{1000}$

$n = 10000$ (0.00004) $P(A_2 \mid A_1) = \frac{4000}{9999}$; $P(A_2) = \frac{4000}{10000}$

60% red balls & 40% blue balls.

$P(A_2 \mid A_1) \longrightarrow P(A_2)$ as $n \to \infty$

Recap:

- R
- Ω, $\mathcal{F}$, P
- axiomatic definition of probability.
- several properties
- Uniform probability spaces
- conditional probability
- total probability law
- Baye's rule
- independence of events
- setting new probability space from repeating indep. Bernoulli trials.
- sampling with /without replacement.

Tutorial sheet # 1.

Problem : A machine contains 4 components in parallel with 0.1, 0.2, 0.3, 0.4 as their probabilities of failures resp. The machine fails if all the components fail simultaneously. The failure of components is independent. Then what is the probability that the machine once started will not fail??

→ Let A_i : Event that component i fails.

$A_1 \cap A_2 \cap A_3 \cap A_4$: machine fails

$$P(A_1 \cap \dots A_4)^c = 1 - P(A_1 \cap A_2 \cap A_3 \cap A_4)$$ independence.

$$= 1 - P(A_1)\,P(A_2)\,P(A_3)\,P(A_4)$$

$$= 1 - 0.1 \times 0.2 \times 0.3 \times 0.4$$

Problem: A course on Probability is a very famous course and students are allowed to register after the teacher consents. It is observed that 20% times after obtaining the consent, student fails to register for the course. There are only 100 seats available in the class. If teacher consents 102 students, what is the probability that all the students (who try to register) are accommodated in the class??

Ans: $1 - (0.8)^{102} - 102 \times (0.8)^{101} \times 0.2$

Problem : 3 printers

P_1, P_2, P_3

Printer P_1 gets 50% jobs

P_2 30%

P_3 20%

P_1 has 0.15 as its probability of failure,

P_2 0.1

P_3 0.2

If a randomly selected printing job is a failure, what is the probability that it was printed by printer 1??

Problem: If $A_1, A_2 \in \mathcal{F}$ are indep., can we say A_1^c & A_2^c are also indep. ??

Ans: Yes

$$(A_1 \cup A_2)^c = A_1^c \cap A_2^c$$

$$P(A_1 \cup A_2)^c = 1 - P(A_1 \cup A_2)$$

$$\begin{aligned} P(A_1^c \cap A_2^c) &= 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2) \\ &= 1 - P(A_1) - P(A_2) + P(A_1) P(A_2) \\ &= (1 - P(A_1))(1 - P(A_2)) \\ &= P(A_1^c) P(A_2^c) \end{aligned}$$

$\Rightarrow A_1^c$ and A_2^c are indep.

Problem: $A_1, A_2 \in \mathcal{F}$ are indep.

Are A_1, A_2^c also indep??

$p \Rightarrow q$

$\equiv$

$\neg q \Rightarrow \neg p$

Soln:

$A_1 = (A_1 \cap A_2^c) \cup (A_1 \cap A_2)$

$P(A_1) = P(A_1 \cap A_2^c) + P(A_1 \cap A_2)$

$= P(A_1 \cap A_2^c) + P(A_1)P(A_2)$

$P(A_1 \cap A_2^c) = P(A_1) - P(A_1)P(A_2)$

$= P(A_1)(1 - P(A_2))$

$$\boxed{P(A_1 \cap A_2^c) = P(A_1)\,P(A_2^c)}$$

$\Rightarrow A_1$ & A_2^c are indep. $\blacksquare$