ST2131

AY21/22 SEM 2

github/jovyntls

01. COMBINATORIAL ANALYSIS

factorials - $1! = 0! = 1$

N1 - if we know how to count the number of different ways that an event can occur, we will know the probability of the event.

N2 - there are $n!$ different arrangements for n objects.

N3 - there are $\frac{n!}{n_1!\,n_2!\dots n_r!}$ different arrangements of n objects, of which n_1 are alike, n_2 are alike, ..., n_r are alike.

Combinations

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!} = \binom{n-1}{r-1} + \binom{n-1}{r}, \quad 1 \le r \le n$$

N5 - **Binomial Theorem**: $(x+y)^n = \sum\limits_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

Multinomial Coefficients

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1!\,n_2!\dots n_r!}$$

N6 - represents the number of possible divisions of n distinct objects into r distinct groups of respective sizes $n_1, n_2, \dots, n_3$, where $n_1 + n_2 + \dots + n_r = n$

N7 - **Multinomial Theorem**: $(x_1 + x_2 + \dots + x_r)^n = \sum\limits_{(n_1,\dots,n_r):n_1+n_2+\dots+n_r=n} \frac{n!}{n_1!\,n_2!\dots n_r!} x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}$

Number of Integer Solutions of Equations

N8 - there are $\binom{n-1}{r-1}$ distinct *positive* integer-valued vectors $(x_1, x_2, \dots, x_r)$ satisfying

$x_1 + x_2 + \dots + x_r = n, \quad x_i > 0, \quad i = 1, 2, \dots, r$

N9 - there are $\binom{n+r-1}{r-1}$ distinct *non-negative* integer-valued vectors $(x_1, x_2, \dots, x_r)$ satisfying $x_1 + x_2 + \dots + x_r = n$

Proof. let $y_k = x_k + 1 \Rightarrow y_1 + y_2 + \dots + y_r = n + r$

02. AXIOMS OF PROBABILITY

DeMorgan's Laws:

$$\left(\bigcup_{i=1}^{n} \mathbf{E_i}\right)^{\mathbf{c}} = \bigcap_{i=1}^{n} \mathbf{E_i^c} \quad \text{and} \quad \left(\bigcap_{i=1}^{n} \mathbf{E_i}\right)^{\mathbf{c}} = \bigcup_{i=1}^{n} \mathbf{E_i^c}$$

Axioms of Probability

definition 1: relative frequency

$P(E) = \lim_{n\to\infty} \frac{n(E)}{n}$. problems: (1) $\frac{n(E)}{n}$ may not converge when $n \to \infty$. (2) $\frac{n(E)}{n}$ may not converge to the same value if the experiment is repeated.

Axioms (definition 2)

For each event E of the sample space S, we assume that a number $P(E)$ is defined and satisfies the following 3 axioms:

1. $0 \le P(E) \le 1$ 2. $P(S) = 1$

3. For mutually exclusive events, $P(\bigcup\limits_{i=1}^{\infty} E_i) = \sum\limits_{i=1}^{\infty} P(E_i)$.

mutually exclusive $\to E_i E_j = \emptyset$ when $i \ne j$

Simple Propositions

N1 - $P(\emptyset) = 0$

N6 - **probability function** $\iff$ it satisfies the 3 axioms.

N8 - if $E \subset F$, then $P(E) \le P(F)$

N10 - Inclusion-Exclusion identity where $n = 3$

$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)$

N11 - **Inclusion-Exclusion identity** -

$P(E_1 \cup E_2 \cup \dots \cup E_n) = \sum\limits_{i=1}^{n} P(E_i) - \sum\limits_{i_1<i_2} P(E_{i_1} E_{i_2}) + \dots + (-1)^{r+1} \sum\limits_{i_1<\dots<i_r} P(E_{i_1} \dots E_{i_r}) + \dots + (-1)^{n+1} P(E_1 \dots E_n)$

(i) $P(\bigcup_{i=1}^n E_i) \le \sum\limits_{i=1}^{n} P(E_i)$

(ii) $P(\bigcup_{i=1}^n E_i) \ge \sum\limits_{i=1}^{n} P(E_i) - \sum\limits_{j<i} P(E_i E_j)$

Sample Space w/ Equally Likely Outcomes

Consider $S = \{e_1, e_2, \dots, e_n\}$. Then $P(\{e_i\}) = \frac{1}{n}$ or $P(\{e_1\}) = P(\{e_2\}) = \dots = P(\{e_n\}) = \frac{1}{n}$

N1 - for any event E,

$P(E) = \frac{\text{\# of outcomes in } E}{\text{\# of outcomes in } S} = \frac{\text{\# of outcomes in } E}{n}$

increasing sequence of events $\{E_n, n \ge 1\} \to E_1 \subset E_2 \subset \dots \subset E_n \subset \dots$

decreasing sequence of events $\{E_n, n \ge 1\} \to E_1 \supset E_2 \supset \dots \supset E_n \supset \dots$

increasing: $\lim\limits_{n\to\infty} E_n = \bigcup\limits_{i=1}^{\infty} E_i$, decreasing: $\lim\limits_{n\to\infty} E_n = \bigcap\limits_{i=1}^{\infty} E_i$

N2 - for both *increasing* and *decreasing* sequence,

$\lim\limits_{n\to\infty} P(E_n) = P(\lim\limits_{n\to\infty} E_n)$

03. CONDITIONAL PROBABILITY AND INDEPENDENCE

Conditional Probability

if $P(F) > 0$, then $P(E|F) = \frac{P(E \cap F)}{P(F)}$

multiplication rule: $P(E_1 \dots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)\dots P(E_n|E_1E_2\dots E_{n-1})$

N3 - **axioms of probability** apply to conditional probability

1. $0 \le P(E|F) \le 1$
2. $P(S|F) = 1$ where S is the sample space
3. If E_i ($i \in \mathbb{Z}_{\ge 1}$) are mutually exclusive, then $P(\bigcup\limits_{1}^{\infty} E_i|F) = \sum\limits_{1}^{\infty} P(E_i|F)$

N4 - If we define $Q(E) = P(E|F)$, then all previously proven results apply.

- $P(E_1 \cup E_2|F) = P(E_1|F) + P(E_2|F) - P(E_1E_2|F)$

Total Probability & Bayes' Theorem

conditioning formula -

$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$

$P(F|E) = \frac{P(EF)}{P(E)} = \frac{P(F) \cdot P(E|F)}{P(E)}$

$P(F^c|E) = \frac{P(EF^c)}{P(E)} = \frac{P(F^c) \cdot P(E|F^c)}{P(E)}$

Total Probability

theorem of total probability - Suppose $F_1, F_2, \dots, F_n$ are mutually exclusive events such that $\bigcup\limits_{i=1}^{n} F_i = S$, then

$$P(E) = \sum_{i=1}^{n} P(EF_i) = \sum_{i=1}^{n} P(F_i)P(E|F_i)$$

Bayes Theorem

$$P(F_j|E) = \frac{P(EF_j)}{P(E)} = \frac{P(F_j)P(E|F_j)}{\sum_{i=1}^{n} P(F_i)P(E|F_i)}$$

application of bayes' theorem

$$P(B_1 \mid A) = \frac{P(A|B_1) \cdot P(B_1)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)}$$

Let A be the event that the person test positive for a disease. B_1: the person has the disease. B_2: does not have.

true positives: $P(B_1 \mid A)$	false negatives: $P(\bar{A} \mid B_1)$
false positives: $P(A \mid B_2)$	true negatives: $P(\bar{A} \mid B_2)$

Independent Events

N1 - $E \perp F \iff P(EF) = P(E) \cdot P(F)$

N2 - $E \perp F \iff P(E|F) = P(E)$

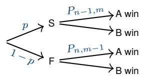
N3 - $E \perp F \iff E \perp F^c$

N4 - if E, F, G are independent, then E will be independent of any event formed from F and G. (e.g. $F \cup G$)

N6 - $(E \perp F) \wedge (E \perp G) \not\Rightarrow E \perp FG$

N7 - For independent trials with probability p of success, probability of m successes before n failures, for $m, n \ge 1$,

method 1



method 2

$P_{n,m} = \sum\limits_{k=n}^{m+n-1} \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}$

$= P(\ge n \text{ successes in } m+n-1 \text{ trials})$

04. RANDOM VARIABLES

Types of Random Variables

- **Bernoulli r.v.** $\to p(x) = \begin{cases} p, & x = 1, \text{ ('success')} \\ 1-p, & x = 0 \quad \text{('failure')} \end{cases}$
- **Binomial r.v.** $\to Y = X_1 + X_2 + \dots + X_n$ where $X_1, X_2, \dots, X_n$ are independent Bernoulli r.v.'s.
  - $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
  - $P(k$ successes from n independent trials each with probability p of success$)$

    $E(Y) = np, \quad Var(Y) = np(1-p)$
- **Negative Binomial** $\to X =$ # trials until k successes
  - $E[X] = k/p$
- **Geometric** $\to X =$ number of trials until a success
  - $P(X = k) = (1-p)^{k-1} \cdot p$ where $k =$ # trials needed
  - $E[X] = 1/p$
- **Hypergeometric** $\to X =$ number of trials until success, *without replacement* (for m red balls of N balls)
  - $P(X = k) = \binom{m}{k}\binom{N-m}{n-k} / \binom{N}{n}$, $k = 0, 1, \dots, n$
  - $E[X] = rn/N$

Properties

N1 - if $X \sim \text{Binomial}(n, p)$, and $Y \sim \text{Binomial}(n-1, p)$, then $E(X^k) = np \cdot E[(Y+1)^{k-1}]$

N2 - if $X \sim \text{Binomial}(n, p)$, then for $k \in \mathbb{Z}^+$,

$P(X = k) = \frac{(n-k+1)p}{k(1-p)} \cdot P(X = k-1)$

Coupon Collector Problem

There are N distinct types of coupons. T denotes the number of coupons needed to be collected for a complete set. What is $P(T = n)$? Ans: $P(T > n) = \sum\limits_{i=1}^{N-1} \binom{N}{i} \left(\frac{N-i}{N}\right)^n (-1)^{i+1}$

Probability Mass Function

pmf of X (*discrete*) $\to p(a) = P(X = a)$

- $\sum_{i=1}^{\infty} p(x_i) = 1$

Cumulative Distribution Function

- **cdf** of a r.v. $X \to$ the function F defined by

  $F(x) = P(X \le x), \quad -\infty < x < \infty$
  - $F(x)$ is defined on the entire real line.

pmf,

a	1	2	4
$p(a)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

$F(a) = \sum\limits_{x \le a} p(x)$ for all

cdf, $F(a) = \begin{cases} 0, & a < 1 \\ 1/2, & 1 \le a < 2 \\ 3/4, & 2 \le a < 4 \\ 1, & a \ge 4 \end{cases}$

Expected Value, μ

discrete: $E(X) = \sum_x x \cdot p(x)$

continuous: $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

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

N1 - if a and b are constants, then $E(aX + b) = aE(X) + b$

N3 - for a non-negative r.v. Y, $E(Y) = \int_0^{\infty} P(Y > y)\,dy$

- I is an indicator variable for event A if $I = \begin{cases} 1, \text{ if } A \text{ occurs} \\ 0, \text{ if } A^c \text{ occurs} \end{cases}$.

  then $E(I) = P(A)$.

2 methods for finding expectation of f(x)

1. using pmf of Y: let $Y = f(X)$. Find X for each Y.
2. using pmf of X: $E[f(x)] = \sum_x f(x)p(x)$

Variance

For X with mean $\mu = E[X]$, the **variance** of X is

$Var(X) = E[(X-\mu)^2] = E(x^2) - [E(x)]^2$

- $Var(aX + b) = a^2 Var(x)$
- $Var(X) = \sum_{x_i} (x_i - \mu)^2 \cdot p(x_i)$ (deviation · weight)

Poisson Random Variable

X is a **Poisson r.v.** with parameter λ if for some $\lambda > 0$,

$$P(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

$$E(X) = \lambda, \quad Var(X) = \lambda$$

- **Poisson Approximation of Binomial** - if $X \sim \text{Binomial}(n, p)$, where n is large and p is small, then $X \dot\sim \text{Poisson}(\lambda)$ where $\lambda = np$. (✓ *weak* dependence is ok)

Poisson distribution as random events

Let $N(t)$ be the number of events in time interval $[0, t]$.

N1 - If the 3 assumptions are true, then $N(t) \sim \text{Poisson}(\lambda t)$.

N2 - If λ is the *rate of occurrences* of events per unit time, then the number of occurrences in an interval of length t has a Poisson distribution with mean λt.

$P(N(t) = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$, for $k \in \mathbb{Z}_{\ge 0}$

o(h) notation

$o(h)$ stands for any function $f(h)$ such that $\lim\limits_{h\to 0} \frac{f(h)}{h} = 0$

- $o(h) + o(h) = o(h)$
- $\frac{\lambda t}{n} + o(\frac{t}{n}) \doteq \frac{\lambda t}{n}$ for large n

Expected Value of sum of r.v.

For a r.v. X, let $X(s)$ denote the value of X when $s \in \mathcal{S}$

N1 - $E(x) = \sum\limits_i x_i P(X = x_i) = \sum\limits_{s \in \mathcal{S}} X(s)p(s)$ where $\mathcal{S}_i = \{s : X(s) = x_i\}$

N2 - $E(\sum\limits_{i=1}^{n}) = \sum\limits_{i=1}^{n} E(X_i)$ for r.v. $X_1, X_2, \dots, X_n$