Sets, Permutation, Stirling's Approximation

Tuesday, April 4, 2023 10:05 AM

Set: Collection of distinct elements = {a,b,c,…}
- Order doesn't matter
- N-set: set of size n
- 0-set: empty set

Permutations of n-set: n!
- Number of ordering of its elements
- 0! = 1 (1 way to order nothing)

Stirling's Approximation: $n! \approx \sqrt{2\pi n}(n/e)^n$

Partial Permutations, Combinations

Thursday, April 6, 2023 9:32 AM

Partial permutation: Ordering of some elements in a set

Notation: k-permutation of n-set = ordering of k elements in a set of size n

- $P(n,k) = \frac{n!}{(n-k)!}$

Combinations: k-subset of some size of a set

- $C(n,k) = \binom{n}{k} = \frac{P(n,k)}{k!} = \frac{n!}{k!\,(n-k)!}$

Probability, Events

Tuesday, April 11, 2023 9:34 AM

Sample space: set of all possible outcomes of an experiment
- Continuous: uncountably infinite
- Discrete: finite or countable infinite

Random Outcome: denoted by uppercase variable
- Before experiment: outcome is unknown
- After experiment: outcome is known
- Outcome called observation

Probability of an outcome: P(X) is the fraction of times X will occur when experiment is repeated
- $P(x) = \lim_{n\to\infty} \frac{n_x}{n}$
  Since $0 \le n_x \le n$ then $0 \le P(x) \le 1$
  And $\sum_{x\in\Omega} P(x) = 1$
- Maps each outcome to a probability
- Probability space: set of all probabilities for each outcome, must sum to 1

Uniform Probability Spaces: all outcomes are equally likely

$$P(x) = \frac{1}{|\Omega|}$$

Non uniform probability spaces: not all outcomes are equally likely

Event E is a subset of the sample space, occurs if $x \in E$
Probability of subset (event):

$$P(E = \{a, b, c, \ldots\}) = \sum_{x\in E} P(x)$$

Multiple Events, Repeated Experiments

Thursday, April 13, 2023 9:34 AM

Def: Sets are disjoint if they do not share any elements. Events are mutually exclusive if they are disjoint

Given events A, B:

If $A \subseteq B \rightarrow P(A) \leq P(B)$

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

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

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

Repeated Experiments: can represent as one experiment where outcomes are tuples or sets of each experiment outcomes

Independent experiments: experiment outcomes are independent
- Sample space: Ω^n

Dependent experiments: each experiment influences the next experiment
- Sample space: $\Omega^{\underline{n}}$

Probability Axioms

Tuesday, April 18, 2023 10:05 AM

Non-negativity: $P(\forall X \in \Omega) \geq 0$
Unitarity: $P(\Omega) = 1 \rightarrow P(\forall X \in \Omega) \leq 1$
Additivity: $A, B\ disjoint \rightarrow P(A \cup B) = P(A) + P(B)$

Complement rule: $P(A^c) = 1 - P(A)$
Subtraction: $P(A - B) = P(A) - P(A \cap B)$
Inclusion-exclusion (2 sets): $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Inclusion-exclusion (n sets): $P(S_1 \cup \cdots \cup S_n) = P(S_1) + \cdots P(S_n) - P(S_1 \cap S_2) - \cdots P(S_{n-1} \cap S_n) + P(S_1 \cap S_2 \cap s_3) + \cdots P(S_{n-2} \cap S_{n-1} \cap S_n) - \cdots$
Alternate adding and subtracting n-wise intersections of all sets
Null rule: $P(\emptyset) = 0$

Conditional Probability, Correlation, Independence

Thursday, April 20, 2023 10:32 AM

Given events E, F - probability that F occurs given E occurs

For uniform spaces: $P(F|E) = \frac{|E \cap F|}{|E|}$

In general: $P(F|E) = \frac{P(E \cap F)}{P(E)}$

Properties:

$P(B\ |A) \geq 0$
$P(\Omega|A) = 1$
$P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$

Correlation:

Positive: $P(F|E) \geq P(F) \rightarrow P(E|F) \geq P(E)$

Negative: $P(F|E) \leq P(F) \rightarrow P(E|F) \leq P(E)$

Independent Events: Events which do not affect the other's probability are called statistically independent

Given independent events E, F: $P(F|E) = P(F) \rightarrow P(E|F) = P(E)$

Thus: $P(E \cap F) = P(E) * P(F)$

Mutual independence: Given events E, F, G then $P(EFG) = P(E) * P(F) * P(G)$

If E and F are independent, then E and F^c are independent

Product Rule:

$P(E \cap F) = P(E) * P(F|E)$
$P(E \cap F \cap G) = P(E) * P(F|E) * P(G|E \cap F)$

Law of Total Probability

Thursday, April 27, 2023 10:31 AM

Def: Let $E = \{E_1, E_2, \ldots, E_n\}$ be partitions of Ω

$$P(F) = \sum_{e \in E} P(F \cap e) = P(e) * P(F|e)$$

Bayes' Rule

Tuesday, May 2, 2023 9:32 AM

Given $P(F|E)$, we want to find $P(E|F)$:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E) * P(F|E)}{P(F)}$$

Random Variables, CDF, PDF

Tuesday, May 2, 2023 10:10 AM

PMF is a mapping from the sample space to their probabilities $P : \Omega \to R$

Discrete: maps to finite or countably infinite set
Continuous: maps to uncountably infinite set

Idea: we can use random variables in functions

$Y = X + b$: $P(Y = y) = P(X + b = y) = P(X = y - b)$
$Y = X * b$: $P(Y = y) = P(b * X = y) = P\left(X = \frac{y}{b}\right)$
$Y = f(X)$: $P(Y = y) = P(f(X) = y) = P(X \in f^{-1}(y))$

Cumulative Distribution Function:

$$F(x) = P(X \leq x) = \sum_{u \leq x} P(u)$$

$x \leq y \to F(x) \leq F(y)$
$\lim_{x \to -\infty} F(x) = 0$
$\lim_{x \to \infty} F(x) = 1$

$P(X > x) = 1 - F(x)$
$P(a < X \leq b) = F(b) - F(a)$

Given CDF, we can find the PMF:

$P(i) = F(i) - F(i - 1)$

PDF: $P(x) = \frac{d}{dx} F(x)$

Expected Value

Thursday, May 4, 2023 10:03 AM

Range:

$$x_{min} = \min\{x \in \Omega : P(x) > 0\}$$
$$x_{max} = \max\{x \in \Omega : P(x) > 0\}$$

Expected Value: Given random variable X

$$E(X) = \sum_{x \in X} P(x) * x$$
$$x_{min} \leq E(X) \leq x_{max}$$
$$E(X = c) = c \rightarrow E\big(E(X)\big) = E(X)$$

For uniform variables: expected value is arithmetic mean
For symmetric variables around s: expected value is s
For non-negative variables:

Modified Variables: Given Y = g(X)

$$E(Y) = \sum P(y) * y = \sum P(x) * g(x)$$

Linearity of Expectations:

$$E(X + b) = E(X) + b$$
$$E(c * X) = c * E(X)$$

Conditional Expectation, Total Expectation

Thursday, May 11, 2023 9:50 AM

Given events X and A:

$$E(X|A) = \sum_{x \in A} P(x|A) * x$$

Given events $A_1, A_2, \ldots, A_n$ partitions of Ω

$$E(X) = \sum_{n} P(A_i) * E(X|A_i) = \sum_{n} P(A_i) * \sum_{x \in A_i} x * P(x|A_i)$$

Variance, Standard Deviation

Thursday, May 11, 2023 10:22 AM

Variance: $V(X) = E((X-\mu)^2)$

Standard deviation: $\sigma_X = \sqrt{V(X)}$

Modifications:

$$V(X+b) = V(X)$$
$$V(aX) = a^2 V(X)$$

Simplification: $V(X) = E\left(X^2\right) - E(X)^2$

Indicator Random Variables

Thursday, May 18, 2023 10:18 AM

Def: indicator variable $I_j = \begin{cases} 1 \ if \ X_j = j \\ 0 \ if \ X_j \neq j \end{cases}$

Two Random Variables

Tuesday, May 16, 2023 9:56 AM

Note: $P(x,y) = P(X = x \wedge Y = y)$

Marginals: $P(x) = \sum_y p(x,y)$, $P(y) = \sum_x p(x,y)$

Independence: $P(x,y) = P(x) * P(y)$

Functions on two variables: Given X, Y, and g(x,y)

If g(x,y) = x + y then $P(x+y) = \sum_{u=-\infty}^{\infty} P(u) * P(V = s - u)$ which is the convolution

$$E(g(X,Y)) = \sum_{x,y} g(x,y)p(x,y)$$

$$E(X+Y) = E(X) + E(Y)$$

Note: the expected value of dependent or independent variables are the same

Expected Value of Two Conditional Variables

Tuesday, May 23, 2023 9:32 AM

Expected value of Y given X=x: $E(Y|x)$
Expected value of Y given some X: $E(Y|X)$

Properties:

$$E(X+Y|z) = E(X|z) + E(Y|z)$$
$$E\big(E(X|Y)\big) = E(X)$$

Properties of Two Independent Variables

Tuesday, May 23, 2023 9:51 AM

Given independent X, Y both are equivalent:

$$E(X * Y) = E(X) * E(Y)$$
$$E(X + Y) = E(X) + E(Y)$$
$$V(X + Y) = V(X) + V(Y)$$

Covariance, Correlation, Correlation Coefficient

Tuesday, May 23, 2023 10:08 AM

Covariance: $E\left((x-\mu_x)*(Y-\mu_y)\right) = E(X*Y) - E(X)*E(Y)$

Properties of Covariance:

$Cov(X,X) = V(X)$
$Cov(X,Y) = Cov(Y,X)$
$Cov(X+a,Y) = Cov(X,Y)$
$Cov(aX,Y) = a*Cov(X,Y)$

Correlation:

Positive Correlation: $Cov(X,Y) > 0$
No Correlation: $\mathrm{Cov(X,Y)} = 0$
Negative Correlation: $\mathrm{Cov(X,Y)} < 0$

Note: Independence implies no correlation, no correlated does not imply independence

Correlation Coefficient:

Issue: $Cov(aX,aY) = a^2 Cov(X,Y)$ which may cause dependence on units and scaling

$$r_{X,Y} = \frac{Cov(X,Y)}{\sigma_X * \sigma_Y}$$

Variance of sum of Two Variables: $V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$

Distribution Families

Thursday, May 25, 2023 10:31 AM

Bernoulli: $B_p = \{0, 1\}$

p(1) = p
p(0) = q = 1 - p

V(B) = p*q

N Independent Bernoulli Trials: $P(X_1 \ldots X_n) = p^{n1} q^{n0}$

Binomial: $B_{p,n}(k)$ is probability of k succeses in n independent Bernoulli p

In general: $B_{p,n}(k) = \binom{n}{k} p^k q^{n-k}$
Expected: $E(B_{p,n}) = np$
Variance: $V\left(B_{n,p}\right) = npq$

Poisson: $P_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ where $\lambda = np$ when $n \gg 1 \gg p$

Approximates Binomial for large n and small p

Expected: $E(P_\lambda) = \lambda$
Variance: $V(P_\lambda) = \lambda$

Geometric Distribution: $G_p = pq^{n-1}$ number of bernoulli trials before first success

Expected: $E\left(G_p\right) = \frac{1}{p}$
Variance: $V\left(G_p\right) = \frac{q}{p^2}$

Triangle: $t(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$

Quadratic: $q(x) = \begin{cases} \frac{1}{x^2} & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$

Exponential: $e_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$
Where lambda is the rate, expected # successes per unit

Gaussian/Normal: $N\left(\mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Standard normal: $N(0, 1)$

Properties of Continuous Distributions

Tuesday, June 6, 2023 9:46 AM

Expectation: $E(X) = \int x * f(x)\, dx$

$$E\big(g(x)\big) = \int g(x) * f(x) dx$$

$$E(a * X + b) = a * E(X) + b$$

Variance: $V(X) = \int f(x)\big(x - \mu^2\big)dx = E\big(X^2\big) - E(X)^2$

$$V(X + b) = V(X)$$
$$V(a * X) = a^2 V(X)$$

Functions: given $y = g(x)$, $p(y) = \frac{f(x)}{|g'(x)|}|x = g^{-1}(x)$

If not 1-1: $p(y) = \sum_{x \in g^{-1}(x)} \frac{f(x)}{|g'(x)|}$