- Combinatorial Analysis -

Distinct Objects, Distinct Boxes

The number of permutations of n objects of which n_1 are of one kind, n_2 are of a second kind, ..., and n_k are of a kth kind is

$$\frac{n!}{n_1!n_2!\cdots n_k!} = \binom{n}{n_1, n_2, \ldots, n_k}$$

Identical Objects, Distinct Boxes

The number of ways of distributing n identical objects into r distinct boxes is the solution to the equation

$$x_1 + x_2 + \cdots + x_r = n$$

where x_i is the number of objects in the ith box. The number of solutions is

$$\binom{n+r-1}{r-1}$$

When we require that each box contain at least one object, the number of solutions is

$$\binom{n-1}{r-1}$$

The proof can also be achieved by using the stars and bars method.

Distinct Objects, Identical Boxes

Whenever there are two or more identical boxes, remember to divide by the number of ways the boxes can be permuted. The number of ways of distributing n distinct objects into r identical boxes is

$$\binom{n+r-1}{r-1}\frac{1}{r!}$$

Multinomial Coefficients

The expansion of $(x_1 + x_2 + \cdots + x_k)^n$ is

$$\sum_{n_1+n_2+\cdots+n_k=n} \frac{n!}{n_1!n_2!\cdots n_k!} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}$$

- Probability -

Concepts

- primitive defination of probability: limiting frequency
- Mutual Exclusivity: pairwise disjoint
- Exhaustivity: union of all events is the sample space

Probability Axioms

1. $P(S) = 1$
2. $P(E) \geq 0$
3. mutually exclusive, straight addition

Inclusion-Exclusion Principle

$$P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} (-1)^{i+1} [\sum_{1<j_1<\cdots<j_i\leq n} P(\bigcap_{k=1}^{i} E_{j_k})]$$

Bayes' Theorem

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$

- sensitivity = probability of a positive test given that the patient has the disease = $\frac{TP}{TP+FN}$
- specitivity = probability of a negative test given that the patient does not have the disease = $\frac{TN}{TN+FP}$
- prevalence = probability of having the disease = $\frac{TP+FN}{TP+FN+TN+FP}$

- Random Variables -

Over a sample space S, a random variable X is a function with the map:

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

or, if X is discrete, the map:

$$X : S \to \mathbb{Z}$$

For any proofs invoving random variables, remember to use the definition of the random variable and probability of random variables.

Bernoulli Distribution

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

We say $X \sim Be(p)$, and

$$E(X) = p,\ Var(X) = p(1-p)$$

Binomial Distribution

If x is the number of successes in n independent Bernoulli trials, then the probability mass function of x is:

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

We say $X \sim Bin(n, p)$, and

$$E(X) = np,\ Var(X) = np(1-p)$$

Geometric Distribution

Define the random variable X as the number of trials until the first success. Then the probability mass function of X is:

$$P(X = x) = (1-p)^{x-1} p$$

We say $X \sim Geo(p)$, and

$$E(X) = \frac{1}{p},\ Var(X) = \frac{1-p}{p^2}$$

negative binomial distribution

Define the random variable X as the number of trials until the rth success. Then the probability mass function of X is:

$$P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

We say $X \sim NB(r, p)$, and

$$E(X) = \frac{r}{p},\ Var(X) = \frac{r(1-p)}{p^2}$$

Hypergeometric Distribution

If X is the number of successes in n draws without replacement from a population of N objects of which K are successes, then the probability mass function of X is:

$$P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

We say $X \sim HGeom(N, K, n)$, and

$$E(X) = \frac{nK}{N},\ Var(X) = \frac{nK(N-K)(N-n)}{N^2(N-1)}$$

Poisson Distribution

If X is the number of events in a fixed interval of time or space, then the probability mass function of X is:

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

We say $X \sim Pois(\lambda)$, and

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

Poison random variables can be used as approximations for binomial random variables when n is large and p is small enough so that $\lambda = np$ is moderate(usually if $n > 20$ and $np < 15$).