Probabilities and Statistics I - Test I: Formulae

Counting

Sample Space (S) = all possible outcomes of a statistical experiment.
Element = each outcome/member of S. $S_0 = s_1, s_2, \ldots s_n \mid s_i \in S_0$
Event = Subspace of a sample space S; s.t. an experiment is an event that generates a data set.

Multiplication Rule: $n(A_1) = n_1, n(A_2) = n_2, \ldots, n(A_K) = n_k \implies n(A, B, \ldots, K) = n_1 n_2 \ldots n_k$

Permutation: An arrangement of all/parts of a set of objects/arranging r objects from a total set of n objects.

$$^{n}P_r = \frac{n!}{(n-r)!} = n!;\ n = r$$

Circular permutation: $^{n}P_r = (n-1)!$.
Permutations for partitioning applications/indistinguishable elements (n_i are number of elements in a subset/partition):

$$^{n}P_r = \binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}$$

Combinations: Selection of r elements from a total set of n elements (two-partition permutation).

$$^{n}C_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Probability of an Event

$$P(S) = P(\phi) + P(S) = \sum_i P(s_i) = P(A) + P(A') = 1;\quad P(A) = \frac{n}{N}$$

Additive Rules:

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

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

$$P\left(\bigcup_{i-1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i;j|i<j} P(A_i \cap A_j) + \sum_{i,j,k|i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

Conditional Probabilities:

$$P(B|A) = \frac{P(A \cap B)}{P(A)};\ P(A) > 0\ or\ *P(B|A) = P(B)$$

* mutually exclusive events

$$\therefore P(A \cap B) = P(A)P(B|A); \; \therefore *P(A \cap B) = P(A)P(B)$$

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1})$$

$$*P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2) \ldots P(A_k)$$

Bayes' Theorem:

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

$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(B_r)P(B_r|A)}{\sum_i P(B_i)P(B_i|A))}$$

Random Variables

A random variable is a function that associates a real number with each element in S.
Discrete Sample Space: S containing a finite number/countably unending sequence of possibilities.
Continuous Sample Space: S containing an infinite number of continuously sequential possibilities.

Probability distribution/mass function (PDF/PMF): A function whose value at any given sample in the sample space can be interpreted as providing a relative likelihood that the value of the random variable would be equal to that sample.

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$ or $\int_{-\infty}^{\infty} dx f(x) = 1$
3. $P(X = x) = f(x)$ or $P(a < X < b) = \int_a^b dx f(x)$

Cumulative distribution function (CDF):

$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$ or $F(x) = \int_{-\infty}^{x} dt f(t); \quad x \in [-\infty, \infty]$
$P(a < X < b) = \int_a^b dx f(x) = \int_{-\infty}^{b} dx f(x) - \int_{-\infty}^{a} dx f(x) = F(b) - F(a)$

* mutually exclusive events