Probability

Notation

Intersection of A and B:	$A \cap B$
Intersection of A and B:	$A \cap B$
Sample space:	$A \cup \overline{A} = S$
Disjoint example:	$A \cap \overline{A} = \emptyset$

Useful Laws and Theorems Derived from Set Notation

De Morgan's laws:	$\overline{A \cup B} = \overline{A} \cap \overline{B}$
Addition laws for probabilities:	$P(A) + P(B) - P(A \cap B) = P(A \cup B)$
Total probability law:	$\Sigma_i P(A_i \cap B) = P(B)$ *where A_i exhausts the sample space*
Conditional probability:	$\frac{P(A \cap B)}{P(B)} = P(A\|B)$
Bayes' theorem:	$P(A\|B)P(B) = P(B\|A)P(A)$

Note: $P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A})$*, which gives the alternative form* $P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|\overline{A})P(\overline{A})}$

Permutations and Combination

Permutations of k objects from n: ${}^nP_k = \frac{n!}{(n-k)!}$

Combinations of k objects from n: ${}^nC_k = \frac{n!}{k!(n-k)!}$

Otherwise known as the binomial coefficient

Distribution Basics

Moment of distribution X:	$\mu_k = E[X^k] = \int x^k f(x) dx$
Variance of distribution X:	$V[X] = E[(X-\mu)^2] = E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - \mu^2$
Central moment of distribution X:	$\nu_k = E[(X-\mu)^k] = \int (x-\mu)^k f(x) dx$
Normalised central moment:	$\gamma_k = \frac{\nu_k}{\nu_k^{\frac{k}{2}}} = \frac{\nu_k}{\sigma^k}$

where γ_3 is the skewness and γ_4 is the kurtois of a distribution

Random Variables

Discrete Random Variables

Define X to be a discrete RV which takes only the values $x_i = 1, 2, ..., n$, then Y is discretely defined to be $y_i = Y(x_i)$

Probability function of Y: $$g(y) = \begin{cases} \Sigma_j f(x_j) & \text{if } y = y_i \\ 0 & \text{otherwise} \end{cases}$$

where the sum extends over j for which $y_i = Y(x_j)$

If $Y(x)$ has a single-valued inverse then we have the closed form expression for $g(y)$

Probability function of Y: $$g(y) = \begin{cases} f(x(y_i)) & \text{if } y = y_i \\ 0 & \text{otherwise} \end{cases}$$

Continuous Random Variables

If X is a continuous RV then so to is $Y(X)$

Probability of Y for $y \to y + dy$: $g(y)dy = \int_{dS} f(x)dx$

If $X(Y)$ is single-valued we have: $g(y)dy = |\int_{x(y)}^{x(y+dy)} f(x')dx'| = \int_{x(y)}^{x(y)+|\frac{dx}{dy}|dy} f(x')dx' = f(x(y))|\frac{dx}{dy}|dy$

Functions of Several Random Variables

Probability function of $Z(X,Y)$: $p(z) = \sum_{i,j} f(x_i, y_i)$

where the sum extends over i, j such that $z = Z(x_i, y_i)$

For the continuous case: $p(z)dz = \int \int_{dS} f(x,y)dxdy$

where dS is the infinitesimal area between $Z(x,y) = z$ and $Z(x,y) = z + dz$

Expectation of Z: $E[Z] = \int zp(z)dz = \int \int Z(X,Y)f(x,y)dxdy$

Taylor expansion of $Z(X,Y)$: $Z \approx Z(\mu_x, \mu_y) + (X - \mu_x)(\frac{\partial Z}{\partial x})_{x=\mu_x} + (Y - \mu_y)(\frac{\partial Z}{\partial y})_{y=\mu_y}$

where the expansion is taken around the means of x and y

Expectation of Taylor expansion: $E[Z] \approx Z(\mu_x, \mu_y) + (E[X] - \mu_x)(\frac{\partial Z}{\partial x})_{x=\mu_x} + (E[Y] - \mu_y)(\frac{\partial Z}{\partial y})_{y=\mu_y}$

$E[Z] \approx Z(\mu_x, \mu_y)$

Variance of $Z(X,Y)$: $V[Z] = \int (z - \mu_z)^2 p(z)dz = \int \int [Z - \mu_z]^2 f(x,y)dxdy$

Hence if x and y are two independent random variables, ie $f(x,y) = g(x)h(y)$

setting $Z = aX + BY + c$ yields: $V[Z] = a^2V[X] + b^2V[X]$

Variance can be approximated as: $V[Z] \approx V[X](\frac{\partial Z}{\partial x})^2_{x=\mu_x} + V[Y](\frac{\partial Z}{\partial y})^2_{y=\mu_y}$

Generating Functions

Probability Generating Functions

The probability of X taking value n can be written as f_n, such that: $\sum_{n=0}^{\infty} f_n = 1$

Probability generating function: $\Phi_X(t) = \sum_{n=0}^{\infty} f_n t^n = E[t^X]$

Sums of Random Variables

Define the sum as: $S_2 = X + Y$

The PGF of the sum has coefficients for t^n that are the probability that $X + Y = n$, this is equivalent to $X = r$ and $Y = n - r$ for $0 \leq$ r $\leq$ n.

Hence we have: $P(S_2 = n) = \sum_{r=0}^{n} P(X = r)P(Y = n - r)$

This leads to the PGF: $\Phi_{S_2}(t) = \sum_{r=0}^{\infty} \sum_{n=r}^{\infty} P(X = r)t^r P(Y = n - r)t^{n-r}$

Setting $n = r + s$ yields: $\Phi_{S_2}(t) = \sum_{r=0}^{\infty} P(X = r)t^r) \sum_{s=0}^{\infty} P(Y = s)t^s = \Phi_X(t)\Phi_Y(t)$

This result can be seen from: $E[t^{X+Y}] = E[t^X]E[t^Y]$

The sum of a variable length of random numbers, ie $S_N = X_1 + X_2 + ... + X_N$, where N is defined to have $P(N = n) = h_n$ with PGF $\chi_N(t) = \sum_n h_n t^n$, can be predicted using:

Probability that $S_N = k$ $\zeta_k = \sum_{n=0}^{\infty} P(N = n)P(S_N = k) = \sum_{n=0}^{\infty} h_n$ x coeff of $t^k in[\Phi_X(t)]^n$