4.2 Combinatorial:

Permutation: A rearrangement of objects into distinct sequence (i.e., order matters). $P_n^m = \frac{m!}{(m-n)!}$

Property: There are $\frac{n!}{n_1!n_2!\dots n_r!}$ different permutations of n objects, of which n_1 are alike, n_2 are alike, $\cdots$, n_r are alike. with alike

Combination: An unordered collection of objects (i.e., order doesn't matter).

Property: There are $\binom{n}{r} = \frac{n!}{(n-r)!r!}$ different combinations of n distinct objects taken r at a time.

Binomial theorem: $(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$ $(a+b)^2 + (a-b)^2 = 2\times(a^2+b^2)$

(a+b) ^n + (a-b) ^n always only left with even number powers (power term is 2)

Inclusion-Exclusion Principle: $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1E_2)$

$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - P(E_1E_3) - P(E_2E_3) + P(E_1E_2E_3)$

and more generally,

$$P(E_1 \cup E_2 \cup \dots \cup E_N) = \sum_{i=1}^{N} P(E_i) - \sum_{i_1<i_2} P(E_{i_1}E_{i_2}) + \cdots + (-1)^{r+1} \sum_{i_1<i_2<\dots<i_r} P(E_{i_1}E_{i_2}\dots E_{i_r}) + \cdots + (-1)^{N+1}P(E_1E_2\cdots E_N)$$

if $E_1 \sim E_N$ same P.

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + \frac{1\cdot(-1)^{N+1}}{N!}$$

where $\sum_{i_1<i_2<\dots<i_r} P(E_{i_1}E_{i_2}\cdots E_{i_r})$ has $\binom{N}{r}$ terms.

4.3 Conditional probability and Bayes' formula:

Conditional probability $P(A|B)$: If $P(B) > 0$, then $P(A|B) = \frac{P(AB)}{P(B)}$ is the fraction of B outcomes that are also A outcomes.

Multiplication Rule: $P(E_1E_2\cdots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)\cdots P(E_n|E_1\cdots E_{n-1})$.

Law of total probability: for any mutually exclusive events $\{F_i\}$, $i = 1, 2, \cdots, n$, whose union is the entire sample space ($F_i \cap F_j = \Phi, \forall i \neq j$; $\bigcup_{i=1}^{n} F_i = \Omega$), we have

$$P(E) = P(EF_1) + P(EF_2) + \cdots + P(EF_n) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$
$$= P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + \cdots + P(E|F_n)P(F_n)$$

Independent events: $P(EF) = P(E)P(F) \Rightarrow P(EF^C) = P(E)P(F^C)$.

Independence is a symmetric relation: X is independent of Y $\Leftrightarrow$ Y is independent of X.

Bayes' Formula: $P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}$ if F_i, $i = 1, \cdots, n$, are mutually exclusive events whose union is the entire sample space.

if mutually exclusive $P(E) =$

$P(B|A) = \frac{P(A|B) * P(B)}{P(A)} = \frac{P(A|B)\cdot P(B)}{P(A|B)\cdot P(B) + P(A|B^c)P(B^c)}$

law of tot. prob.

As the following examples will demonstrate, not all conditional probability problems have intuitive solutions. Many demand logical analysis instead.

4.5 Expected Value, Variance & Covariance

Expected value, variance and covariance are indispensable in estimating returns and risks of any investments. Naturally, they are a popular test subject in interviews as well. The basic knowledge includes the following: $E(X+Y) = E(X) + E(Y)$

linearity of E

If $E[x_i]$ is finite for all $i = 1, \cdots, n$, then $E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n]$. The relationship holds whether the x_i's are independent of each other or not.

$E(X^2) = Var(X) + (E(X))^2$

If X and Y are independent, then $E[g(X)h(Y)] = E[g(x)]E[h(Y)]$.

Covariance: $Cov(X,Y) = E[(X-E[X])(Y-E[Y])] = E[XY] - E[X]E[Y]$.

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

Correlation: $\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

If X and Y are independent, $Cov(X,Y) = 0$ and $\rho(X,Y) = 0$.[26]

$cov(X,Y) = E(XY) - E(X)\cdot E(Y)$

if indp. $E(X\cdot Y) = E(X)\cdot E(Y)$

General rules of variance and covariance:

$Cov(\sum_{i=1}^{n} a_iX_i, \sum_{j=1}^{m} b_jY_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_ib_jCov(X_i, Y_j)$ $cov(aX, bY) = a\cdot b\, cov(X,Y)$

$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, X_j)$

[26] The reverse is not true. $\rho(X,Y) = 0$ only means X and Y are uncorrelated; they may well be dependent.

linearity of cond' expectation:

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

Conditional expectation and variance: by changing y, E of g(x) changes

For discrete distribution: $E[g(X)|Y=y] = \sum_x g(x)p_{X|Y}(x|y) = \sum g(x)p(X=x|Y=y)$ linearity of E

For continuous distribution: $E[g(X)|Y=y] = \int_{-\infty}^{\infty} g(x)f_{X|Y}(x|y)dx$ addition from cond. P.

Law of total expectation: The EV of the conditioned EV of X given Y is the same as E(X)

$$E[X] = E[E[X|Y]] = \begin{cases} \sum_y E[X|Y=y]p(Y=y), \text{ for discrete } Y \\ \int_{-\infty}^{\infty} E[X|Y=y]f_Y(y)dy, \text{ for continuous } Y \end{cases}$$

This is a random var depends on y

4.6 Order Statistics

Let X be a random variable with cumulative distribution function $F_X(x)$. We can derive the distribution function for the minimum $Y_n = \min(X_1, X_2, \cdots, X_n)$ and for the maximum $Z_n = \max(X_1, X_2, \cdots, X_n)$ of n IID random variables with cdf $F_X(x)$ as

regardless of dist. Need to use cdf F(x)

min: $P(Y_n \geq x) = (P(X \geq x))^n \Rightarrow 1 - F_{Y_n}(x) = (1 - F_X(x))^n \Rightarrow f_{Y_n}(x) = nf_X(x)(1 - F_X(x))^{n-1}$

max: $P(Z_n \leq x) = (P(X \leq x))^n \Rightarrow F_{Z_n}(x) = (F_X(x))^n \Rightarrow f_{Z_n}(x) = nf_X(x)(F_X(x))^{n-1}$