Probability Refresher

Let X and Y be random variables that take on values $x_1 \cdots x_I$ and $y_1 \cdots y_J$, respectively, where x_i and y_j are real numbers.

$Pr(x_i)$ = probability that $X = x_i$

$Pr(y_j)$ = probability that $Y = y_j$

$$\therefore \sum_{i=1}^{I} Pr(x_i) = 1$$

$$\& \sum_{j=1}^{J} Pr(y_j) = 1$$

Expected Value

Definition:

$$E(X) = \sum_{i=1}^{I} x_i \cdot Pr(x_i)$$

$$E(Y) = \sum_{j=1}^{J} y_j \cdot Pr(y_j)$$

Properties:

① For any constant, b,

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

& $$E(b \cdot X) = b \cdot E(X)$$

Special case: using $b = -E(X)$,

$$E[X - E(X)] = 0$$

② $$E(X + Y) = E(X) + E(Y)$$

Independence

Definition: X + Y are independent if $Pr(X, Y) = Pr(X) \cdot Pr(Y)$

Properties: If X + Y are independent, then

① $X - a$ and $Y - b$ are independent, for any constants, a and b.

Special Case: $X - E(X)$ and $Y - E(Y)$ are independent

② $E(X \cdot Y) = E(X) \cdot E(Y)$

Proof of Property ②:

$$E(X \cdot Y)$$

$$= \sum_{ij} x_i \cdot y_j \cdot Pr(x_i, y_j)$$

$$= \sum_{ij} x_i \cdot y_j \cdot Pr(x_i) \cdot Pr(y_j)$$

$$= \sum_{ij} [x_i \cdot Pr(x_i)] \cdot [y_j \cdot Pr(y_j)]$$

$$= \left[\sum_i x_i \cdot Pr(x_i)\right] \cdot \left[\sum_j y_j \cdot Pr(y_j)\right]$$

$$= E(X) \cdot E(Y)$$

qed.

Variance

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

Special Case: If $E(X) = 0$ then

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

Properties: for any constant, b,

① $Var(X+b) = Var(X)$

Special Case: $Var[X - E(X)] = Var(X)$

② $Var(b \cdot X) = b^2 \cdot Var(X)$

Proof of Property ②:

$$Var(b \cdot X)$$

$$= E[bX - E(bX)]^2$$

$$= E[bX - b \cdot E(X)]^2$$

$$= E\left(b[X - E(X)]\right)^2$$

$$= E\left(b^2 \cdot [X - E(X)]^2\right)$$

$$= b^2 \cdot E[X - E(X)]^2$$

$$= b^2 \cdot Var(X)$$

qed.

Lemma 1:

If X + Y are independent

and $E(X) = E(Y) = 0$

then $Var(X+Y) = Var(X) + Var(Y)$

Proof:

$Var(X+Y)$

$= E(X+Y)^2$

$= E(X^2+Y^2+2XY)$

$= E(X^2) + E(Y^2) + E(2XY)$

$= Var(X) + Var(Y) + 2 \cdot E(X) \cdot E(Y)$

$= Var(X) + Var(Y) + 2 \cdot 0 \cdot 0$

$= Var(X) + Var(Y)$

qed.

Theorem 2:

If X & Y are independent,

then $Var(X+Y) = Var(X) + Var(Y)$.

Proof: Let $X' = X - E(X)$ and $Y' = Y - E(Y)$.

$\therefore$ X' and Y' are independent, and

$E(X') = E(Y') = 0$.

Also,

$Var(X+Y)$

$= Var(X'+Y')$

$= Var(X') + Var(Y')$ by Lemma 1

$= Var(X) + Var(Y)$

qed.

Corollary 3:

If $X_1, X_2, \cdots X_N$ are independent random variables, then

$$Var\left(\sum_i X_i\right) = \sum_i Var(X_i)$$

Proof: Similar to Theorem 2.

Corollary 4:

If $X, X_1, X_2, \cdots, X_N$ are independent and identically distributed (iid) random variables, then

$$Var\left(\sum_{i=1}^{N} X_i\right) = N \cdot Var(X)$$

Proof: Because $X, X_1, \cdots X_N$ all have the same distribution,

$$Var(X) = Var(X_1) = \cdots = Var(X_N)$$

$$\therefore Var\left(\sum_{i=1}^{N} X_i\right)$$

$$= \sum_{i=1}^{N} Var(X_i) \qquad \text{by Corollary 3}$$

$$= \sum_{i=1}^{N} Var(X)$$

$$= N \cdot Var(X)$$

q.ed.

Theorem 5:

If $X, X_1, X_2 \cdots X_N$ are independent and identically distributed (iid), and

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

then $$Var(\overline{X}) = \frac{Var(X)}{N}$$

Proof:

$Var(\overline{X})$

$$= Var\left(\frac{1}{N} \sum_{i=1}^{N} X_i\right)$$

$$= \frac{1}{N^2} \cdot Var\left(\sum_{i=1}^{N} X_i\right)$$ by variance property ②

$$= \frac{1}{N^2} \cdot N \cdot Var(X)$$ by Corollary 4

$$= \frac{1}{N} \cdot Var(X)$$

qed.