$E(X) = \mu_X$

1. Basic expectation Rules

1. $E(X) = \mu_X = \sum x p(x)$.
2. $E(a) = a$.
3. $E(aX) = aE(X)$
4. $E(a+X) = a + E(X)$
5. $E(a+bX) = E(a) + bE(X)$
6. $E(X \pm Y) = E(X) \pm E(Y)$
7. $E(XY) = E(X)E(Y)$ X, Y are independent

2. Variance and Covariance expectation Rules.

8. $V(a) = 0$ 常量无变化.
9. $V(a+X) = V(X)$ just a bias.
10. $V(X) = E(X^2) - \mu_X^2$

proof ①: $V(X) = Cov(X,X) = E(X \cdot X) - E(X) \cdot E(X)$
$= E(X^2) - \mu_X^2$

proof ②: $V(X) = E[X - \mu_X]^2$
$= E[X^2] - 2\mu_X E[X] + E\mu_X^2$
$= E[X^2] - \mu_X^2$

11. $Cov(X,Y) = E(XY) - \mu_X \mu_Y$

proof: $Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$
$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$
$= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y]$
$= \underbrace{E[XY]}_{\text{更大}} - E[X]E[Y]$

记: E[X,Y]是直接将所有矩形加起来,
故更大, 由 Cov(X,Y) 和 E[X]E[Y] 两部分组成.

12. $Cov[X,Y] = 0$, $X \perp Y$

proof: $\cdots = E[XY] - E[X]E[Y]$
$= 0$.

即方差运算导致非线性.

13. $V(aX) = a^2 V(X)$
14. $V(X \pm Y) = V(X) + V(Y) + 2\,Cov(X,Y)$

proof: $V(X \pm Y) = Cov(X \pm Y, X \pm Y)$
$= E(X \pm Y)^2 - [E^2(X \pm Y)]$
$= E[X \pm Y]^2 - [E(X) \pm E(Y)]^2$
$= E[X^2 \pm 2XY + Y^2] - [E^2(X) \pm 2E[X]E[Y] + E^2[Y]]$
$= \pm 2E(XY) - (\pm 2E[X]E[Y])$
$+ \underbrace{[E(X^2) - E^2(X)]}_{V(X)} + \underbrace{[E(Y^2) - E^2(Y)]}_{V(Y)}$

$= V(X) + V(Y) \pm 2\,Cov(X,Y)$

15. $V(XY) \neq V(X)V(Y)$

3. Covariance matrix and expectation.

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

例) $\mathrm{Cov}(X) = E\left[(X-u_X)(X-u_X)^T\right]$

proof :

$$\mathrm{Cov}(X) = E\begin{bmatrix} (x_1-u_{x_1})^2 & (x_1-u_{x_1})(x_2-u_{x_2}) & \cdots & (x_1-u_{x_1})(x_k-u_{x_k}) \\ \vdots & & & \\ (x_n-u_{x_k})(x_1-u_{x_1}) & \cdots & & \end{bmatrix}$$

$$= E\left(\begin{bmatrix} (x_1-u_{x_1}) \\ \vdots \\ (x_k-u_{x_k}) \end{bmatrix} \left[(x_1-u_{x_1})\ (x_2-u_{x_2}) \cdots (x_k-u_{x_k})\right]\right)$$

$$= E\left[(X-u_X)(X-u_X)^T\right]$$