Appendix A Elementary probability

① symmetry: let A_1 and A_2 are disjoint sets, we have

1) $p(A_1 \cup A_2) = p(A_1) + p(A_2)$ 2) if not, $p(A_1 \cup A_2) = p(A_1) + p(A_2) - p(A_1 \cap A_2)$

② conditioning $p(A|B) = \frac{p(A \cap B)}{p(B)}$

③ if A and B are independent, $p(A \cap B) = p(A)\,p(B)$

④ Expectation: $E(X) = \sum_{m=1}^{M} x_m p_m = \sum_{m=1}^{M} x_m\, p(X = x_m)$

1) if X and Y are independent, $E(XY) = E(X)E(Y)$ but verse are not

2) $E(A) + E(B) = E(A+B)$ 3) $E(aA) = aE(A)$

⑤ Variance: $var(X) = E((X - E(X))^2) = E(X^2) - E^2(X)$

1) $var(aX) = a^2 var(X)$ 2) X, Y independent, $var(X+Y) = var(X) + var(Y)$

3) standard deviction of a random variable: $\sigma = \sqrt{var(X)}$

⑥ Inequalities: $p(X \geq a) \leq \frac{E(X)}{a}$ for $a > 0$ (Markov's inequality)

$p(|X - E(X)| \geq \varepsilon) \leq \frac{var(X)}{\varepsilon^2}$ for $\varepsilon > 0$ (Chebyshev's inequality)

⑦ law of large Numbers

Assume that $X_1, X_2, \dots, X_n$ are independent random variables with the same expected value μ and the same variance σ^2. Define $Y = \sum_{i=1}^{n} X_i / n$, we have

$$E(Y) = E\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{n\mu}{n} = \mu, \quad var(Y) = \frac{1}{n^2} var(X_1 + \dots + X_n) = \frac{\sigma^2}{n}$$

and $p(|Y - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$ (when $n \uparrow$, $p \downarrow$)

⑧ Covariance and Regression

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

if we have X, and want to guess Y, we need to choose a, b to estimate Y by $\hat{Y}$

where $\hat{Y} = a + E(Y) + b(X - E(X))$, all I wanna is minimum $(Y-\hat{Y})^2$

and it equals $E((Y-\hat{Y})^2) = a^2 + Var(Y) + b^2 Var(X) - 2b\,cov(X,Y)$, when $a=0$, $b = \dfrac{cov(X,Y)}{Var(X)}$

$\hat{Y} = E(Y) + \dfrac{cov(X,Y)}{Var(X)}(X - E(X))$ (the Linear Least squares Estimate LLSE)