Chapter 5. Inequalities

Inequalities are useful tools in establishing various properties of statistical inference methods. They may also provide estimates for probabilities with little assumption on probability distributions.

5.1 Probability inequalities

Markov's inequality. Let X be a non-negative r.v. and $EX < \infty$. Then for any $t > 0$, $P(X > t) \leq EX/t$.

An immediate corollary of Markov's inequality: For any r.v. X and any constant $t > 0$,

$$P(|X| > t) \leq \frac{E|X|}{t} \quad \text{provided } E|X| < \infty,$$

$$P(|X| > t) \leq \frac{E\{|X|^k\}}{t^k} \quad \text{provided } E\{|X|^k\} < \infty. \tag{1}$$

The tail-probability $P(|X| > t)$ is a useful measure in insurance and risk management in finance. (1) implies that the more moments X has, the smaller the tail probabilities are.

Proof of Markov's inequality. Since $X > 0$,

$$\begin{aligned} EX &= \int_0^\infty xf(x)dx = \int_0^t xf(x)dx + \int_t^\infty xf(x)dx \\ &\geq \int_t^\infty xf(x)dx \geq t\int_t^\infty f(x)dx = tP(X > t). \end{aligned}$$

Chebyshev's inequality. Suppose a r.v. X have mean μ and variance $\sigma^2 \in (0, \infty)$. Then $P(|X - \mu| \geq t) \leq \sigma^2/t^2$ for any $t > 0$.

Remarks. (i) Chebyshev's inequality follows from (1) with X replaced by $X - \mu$ and $k = 2$.

(ii) Replacing t by $t\sigma$, we have $P(|Z| > t) \leq 1/t^2$, where $Z = (X - \mu)/\sigma$ is a standardization of X.

(iii) For any r.v. X with mean 0 and variance 1, it holds that

$$P(|X| > 2) \leq 1/4, \qquad P(|X| > 3) \leq 1/9.$$

Example 1. We flipped a coin n times with Head occurred k $(< n)$ times. Therefore a natural estimate for the probability $p = P(\text{Head})$ is k/n. What is the error $k/n - p$ in this estimation?

Let $X_i = 1$ if Head occurred in the i-th flip, and 0 otherwise. Then $k = \sum_{i=1}^n X_i$, and $k/n = n^{-1} \sum_{i=1}^n X_i \equiv \bar{X}_n$. Note k, therefore also $\bar{X}_n$, may take different value if we repeat the experiment. Hence it makes sense to quantify the estimation error in probability such as $P(|\bar{X}_n - p| > \epsilon)$ for some small constant $\epsilon > 0$.

Note $E(\bar{X}_n) = n^{-1} \sum E X_i = p$, $\text{Var}(\bar{X}_n) = n^{-2} \sum \text{Var}(X_i) = n^{-1}\text{Var}(X_1)$ $= n^{-1} p(1-p)$. It follows from Chebyshev's inequality that

$$P(|\bar{X}_n - p| > \epsilon) \le \frac{p(1-p)}{n\epsilon^2} \le \frac{1}{4n\epsilon^2}.$$

Let $\epsilon = 0.1$ and $n = 500$, $P(|\bar{X}_n - p| > 0.1) \le 1/(20) = 0.05$.

5.2 Inequalities for expectations

Cauchy-Schwartz inequality. Let $E(X^2) < \infty$ and $E(Y^2) < \infty$. Then $E|XY| \leq \{E(X^2)E(Y^2)\}^{1/2}$.

A function g is *convex* if for any x, y and any $\alpha \in [0, 1]$,

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y).$$

If $g''(x) \geq 0$ for all x, g is convex. Examples of convex functions include $g(x) = x^2$ and $g(x) = e^x$.

A function g is *concave* if $-g$ is convex. Examples of concave functions are $g(x) = -x^2$ and $g(x) = \log(x)$.

Jensen's inequality. If g is convex, $E\{g(X)\} \geq g(EX)$.

From Jensen's inequality, we have $E(X^2) \geq (EX)^2$. If $X \geq 0$, $E(1/X) \geq 1/EX$ and $E(\log X) \leq \log(EX)$.