Contents

Theorem 0.1 (Cauchy-Schwarz) $\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y]^2$.

Theorem 0.2 (Markov's Inequality) If $X \geq 0$, then for all a, $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$.

Theorem 0.3 (Chebyshev's Inequality) Let $X \geq 0$, then

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[X^2]}{\varepsilon^2}.$$

Corollary 0.4 Let $\mu = \mathbb{E}[X]$. Then

$$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Theorem 0.5 (Weak Law of Large Numbers) Let $X_1, ..., X_n$ be IID RVs, mean μ. Let $S_n = \sum X_i$. Then

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \to 0 \text{ as } n \to \infty.$$

i.e. $\frac{S_n}{n}$ tends to μ in probability.

Theorem 0.6 (Strong Law of Large Numbers) $\mathbb{P}\left(\frac{S_n}{n} \to \mu \text{ as } n \to \infty\right) = 1$, i.e. $\frac{S_n}{n} \to \mu$ as $n \to \infty$ almost surely. Strong law implies weak law.

Definition 0.7 Covariance of X and Y is

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Note that $\mathrm{Cov}(X, Y) = 0$ does not imply X, Y independent.

Definition 0.8 Correlation coefficient of X and Y is

$$\mathrm{corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\ \mathrm{Var}(Y)}}$$

It lies in $[-1, 1]$.

Definition 0.9 Marginal distribution of X is

$$\mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, Y = y)$$

Definition 0.10 Conditional expectation of X given Y is

$$\mathbb{E}[X \mid Y = y] = \sum_x x\mathbb{P}(X = x \mid Y = y)$$

Can view $\mathbb{E}[X \mid Y]$ as random variable in Y.

Theorem 0.11 (Tower Property of Conditional Expectation, Law of Total Expectation) $\mathbb{E}_Y[\mathbb{E}_X[X \mid Y]] = \mathbb{E}_X[X]$. Equivalently,

$$\mathbb{E}[X] = \sum_i \mathbb{E}[X \mid A_i]\mathbb{P}(A_i)$$

where $A_1, ..., A_n$ is partition of Ω.

Definition 0.12 Let X be RV on $\mathbb{N}_0$. Probability generating function (pgf) of X is

$$p_X(z) = \mathbb{E}[z^X] = \sum_{r=0}^{\infty} \mathbb{P}(X = r) z^r$$

The pgf of X uniquely determines (via derivatives) its distribution.

Theorem 0.13 (Abel's Lemma) $\mathbb{E}[X] = \lim_{z \to 1} p'(z)$.

Theorem 0.14 $\mathbb{E}[X(X-1)] = \lim_{z \to 1} p''(z)$.

Theorem 0.15 If $X_1, ..., X_n$ independent, then pgf of $X_1 + \dots + X_n$ is $p_{X_1} \dots p_{X_n}$.

Definition 0.16 Moment generating function of X is $m(\theta) = \mathbb{E}\left[e^{\theta X}\right]$.

Definition 0.17 mgf determines distribution of X, provided that $m(\theta) < \infty$ for all θ in some interval containing the origin.

Definition 0.18 The r-th moment of X is $\mathbb{E}[X^r]$.

Theorem 0.19 The r-th moment of X is the coefficient of $\frac{\theta^r}{r!}$ in $m(\theta)$, i.e.

$$\mathbb{E}[X^r] = \frac{\mathrm{d}^n}{\mathrm{d}\theta^n} m(\theta)|_{\theta=0} = m^{(n)}(0)$$

Theorem 0.20 If X, Y, independent, then $m_{X+Y}(\theta) = m_X(\theta) m_Y(\theta)$.

Theorem 0.21 (Central Limit Theorem) Let $X_1, ..., X_n$ be IID RVs, $\mathbb{E}[X_i] = \mu$, $\mathrm{Var}(X_i) = \sigma^2 < \infty$. Let $S_n = X_1 + \dots + X_n$. Then

$$\lim_{n \to \infty} \mathbb{P}\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right) = \Phi(b) - \Phi(a) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \,\mathrm{d}t$$

So $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ converges in distribution, $\underset{D}{\to}$, to the standard normal $N(0,1)$.

Theorem 0.22 (Continuity Theorem) Let $X_1, ..., X_n$ have mgs $m_1(\theta), ..., m_n(\theta)$ where $m_n(\theta) \to m(\theta)$ as $n \to \infty$ pointwise. Then $X_n \underset{D}{\to} Y$ where Y has mgf $m(\theta)$.