Beginning of Jan.14, 2021

Theorem: (1.78) Total Expectation Theorem, Continuous

Let X, Y be continuous random variables and assume $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ be continuous. Then

$$\mathbb{E}X = \int_{-\infty}^{\infty} \mathbb{E}(X \mid Y = y) f_Y(y) \, \mathrm{d}y.$$

Some Useful Inequalities

Theorem: (1.91) Jensen's Inequality

Let $\varphi : \mathbb{R} \to \mathbb{R}$. We say φ is **convex** if for all $x, y \in \mathbb{R}$ and $\lambda \in (0,1)$ we have

$$\varphi(\lambda x + (1-\lambda)y) \leqslant \lambda\varphi(x) + (1-\lambda)\varphi(y).$$

We say φ is **strictly convex** if the above inequality can be replaced by $<$.

Jensen's inequality states that if $\mathbb{E}|X| < \infty$ and $\mathbb{E}|\varphi(X)| < \infty$, and if φ is convex, then

$$\varphi(\mathbb{E}X) \leqslant \mathbb{E}\varphi(X).$$

Theorem: (1.92) Markov's Inequality

For all $t > 0$, we have

$$\mathbb{P}(|X| > t) \leqslant \frac{\mathbb{E}|X|}{t}.$$

Moreover, if $n \geqslant 1$ is a positive integer, then

$$\mathbb{P}(|X| \geqslant t) \leqslant \frac{\mathbb{E}|X|^n}{t^n}.$$

Theorem: (1.97) Chebyshev's Inequality

Using $n = 2$ in Markov's inequality applied to the random variable $X - \mathbb{E}X$, we have

$$\mathbb{P}(|X - \mathbb{E}X| \geqslant t) \leqslant \frac{\mathrm{var}(X)}{t^2}$$

or equivalently

$$\mathbb{P}(|X - \mu| \geqslant t\sigma) \leqslant \frac{1}{t^2}.$$

Proposition: (1.107) Sum & Convolution

Let X, Y be continuous, independent random variables. Then

$$f_{X+Y}(t) = (f_X * f_Y)(t)$$

where $*$ denotes the convolution:

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(s) f_Y(t-s)\,\mathrm{d}s.$$

Proof. We use independence and the fact that PDFs are derivatives of CDFs:

$$\mathbb{P}(X+Y \leqslant t) = \int_{\{x+y\leqslant t\}} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{t-x} f_X(x) f_Y(y)\,\mathrm{d}y\,\mathrm{d}x = \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{t-x} f_Y(y)\,\mathrm{d}y\,\mathrm{d}x,$$

so

$$\begin{aligned} f_{X+Y}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(X+Y \leqslant t) \\ &= \frac{\mathrm{d}t}{\mathrm{d}x}\int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{t-x} f_Y(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} f_X(x) \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{t-x} f_Y(y)\,\mathrm{d}y\,\mathrm{d}x = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x)\,\mathrm{d}x. \end{aligned}$$

□

Of course, we have assumed once again that it is well-defined to differentiate w.r.t the integral.

Chapter 1

Modes of Convergence & the Limit Theorems

1.1 Modes of Convergence

Definition: (2.1) Almost Sure (a.s.) Convergence

We say $\{Y_n\}$ converges to Y **almost surely** if

$$\mathbb{P}(\lim_{n\to\infty} Y_n = Y) = 1$$

or equivalently

$$\mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}) = 1.$$

Definition: (2.2) Convergence in Probability

We say $\{Y_n\}$ converges to Y **in probability** if for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|Y_n - Y| > \epsilon) = 0,$$

or equivalently

$$\lim_{n\to\infty} \mathbb{P}(\{\omega \in \Omega : |Y_n(\omega) - Y(\omega)| > \epsilon\}) = 0.$$

Definition: (2.3) Convergence in Distribution

We say $\{Y_n\}$ converges to Y **in distribution** in distribution if

$$\lim_{n\to\infty} \mathbb{P}(Y_n \leqslant t) = \mathbb{P}(Y \leqslant t)$$

for all $t \in \mathbb{R}$ such that $s \mapsto \mathbb{P}(Y \leqslant s)$ is continuous at $s = t$.

Remark. Since a Gaussian has continuous PDF, the CLT, to be stated right below, is indeed a statement about convergence in distribution.

Definition: (2.4) Convergence in L^p

Let $0 < p \leqslant \infty$. We say that $\{Y_n\}$ converges to Y **in L^p** if $\|Y\|_p < \infty$ and

$$\lim_{n\to\infty} \|Y_n - Y\|_p = 0,$$

where

$$\|Y\|_p := \begin{cases} (\mathbb{E}|Y|^p)^{1/p} & \text{if } 0 < p < \infty \\ \operatorname{ess\,sup}|X| = \inf\{c > 0 : \mathbb{P}(|X| \leqslant c\} = 1) & \text{if } p = \infty. \end{cases}$$

Remark.

$$\text{Convergence in distribution} \Longleftarrow \text{Convergence in probability} \Longleftarrow \begin{cases} \text{a.s. convergence} \\ \text{convergence in } L^p \end{cases}$$

The converses are all false.

1.2 The Limit Theorems

Theorem: (2.10) Weak Law of Large numbers, Weak LLN

Let $X_1, \ldots, X_n$ be i.i.d. (independent identically distributed) and assume that $\mu := \mathbb{E}X_1 < \infty$. Then X_n converges to $\mathbb{E}X_1$ in probability, i.e., for $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) = 0.$$

Theorem: (2.11) Strong Law of Large Numbers, Strong LLN

Let $X_1, \ldots, X_n$ be i.i.d. with $\mu := \mathbb{E}X_1 < \infty$. Then $X_n \to \mu$ almost surely, i.e.,

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{X_1 + \ldots + X_n}{n} = \mu\right) = 1.$$