In-Class Problems — Week 15, Mon

Problem 1. Prove that the Central Limit Theorem implies the Weak Law of Large Numbers. *Hint:* The only properties of $N(y)$ needed in the proof are that $\lim_{y\to\infty} N(y) = 1$ and $\lim_{y\to\infty} N(-y) = 0$.

NOTE: We didn't get to the following problem in class.

Problem 2. To clarify the somewhat subtle difference between the Weak and Strong Laws of Large Numbers, we will construct an example of a sequence $X_1, X_2, \ldots$ of mutually independent random variables that satisfies the Weak Law of Large Numbers, but not the Strong Law. The distribution of X_i will have to depend on i, because otherwise both laws would be satisfied.[1]

In particular, let $X_1, X_2, \ldots$ be a sequence of mutually independent random variables such that $X_1 = 0$, and for each integer $i > 1$,

$$\Pr\{X_i = i\} = \frac{1}{2i\log i}, \quad \Pr\{X_i = -i\} = \frac{1}{2i\log i}, \quad \Pr\{X_i = 0\} = 1 - \frac{1}{i\log i}.$$

Note that $\mu = \mathrm{E}[X_i] = 0$ for all i.

(a) Show that $\mathrm{Var}[S_n] = \Theta(n^2/\log n)$. *Hint:* $n/\log n > i/\log i$ for $2 \le i \le n$.

(b) Show that the sequence $X_1, X_2, \ldots$ satisfies the Weak Law of Large Numbers, *i.e.*, prove that for any $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left\{\left|\frac{S_n}{n}\right| \ge \epsilon\right\} = 0.$$

We now show that the sequence $X_1, X_2, \ldots$ does not satisfy the Strong Law of Large Numbers.

(c) (The first Borel-Cantelli lemma.) Let $A_1, A_2, \ldots$ be any infinite sequence of mutually independent events such that

$$\sum_{i=1}^{\infty} \Pr\{A_i\} = \infty. \tag{1}$$

[1] This problem is adapted from Grinstead & Snell, *Intro. to Probability*, Ch.8, exercise 16, pp314–315, where is credited to David Maslen.

Prove that

$$\Pr\{\text{infinitely many } A_i \text{ occur}\} = 1.$$

Hint: We know that the probability that no A_i with $n \geq i \geq r$ occurs is

$$\leq e^{-\operatorname{E}[T_{r,n}]} \tag{2}$$

where $T_{r,n} ::= \sum_{i=r}^{n} I_{A_i}$ is the number of events A_i with $n \geq i \geq r$ that occur. What happens as $n \to \infty$?

(d) Show that $\sum_{i=1}^{\infty} \Pr\{|X_i| \geq i\}$ diverges. *Hint:* $\int 1/(xlogx)\,dx = \log\log x$.

(e) Conclude that

$$\Pr\left\{\lim_{n\to\infty} \frac{S_n}{n} = \mu\right\} = 0. \tag{3}$$

and hence that the Strong Law of Large Numbers *completely* fails for the sequence $X_1, X_2, \ldots$.

Hint:

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1},$$

so if $\lim_{n\to\infty} S_n/n = 0$, then also $\lim_{n\to\infty} X_n/n = 0$.

A Appendix

The *probability density function (pdf)* for a random variable, R, is the function $f_R : \text{range}(R) \to [0, 1]$ defined by:

$$f_R(x) ::= \Pr\{R = x\}.$$

Random variables $R_1, R_2, \ldots$ are *mutually independent* iff

$$\Pr\left\{\bigcap_i [R_i = x_i]\right\} = \prod_i \Pr\{R_i = x_i\},$$

for all $x_1, x_2, \cdots \in \mathbb{R}$. They are *k-wise independent* iff $\{R_i \mid i \in J\}$ are mutually independent for all subsets $J \subset \mathbb{N}$ with $|J| = k$.

Theorem (Weak Law of Large Numbers). *Let $S_n ::= \sum_{i=1}^{n} X_i$, where $X_1, \ldots, X_n, \ldots$ are pairwise independent variables with the same expectation, μ, and standard deviation, σ. For any $\epsilon > 0$,*

$$\lim_{n\to\infty} \Pr\left\{\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right\} = 0.$$

Theorem (The Strong Law of Large Numbers). *Let $S_n ::= \sum_{i=1}^{n} X_i$ where $X_1, \ldots, X_i, \ldots$ are mutually independent, identically distributed random variables with finite expectation, μ. Then*

$$\Pr\left\{\lim_{n\to\infty} \frac{S_n}{n} = \mu\right\} = 1.$$

Definition. For any random variable, R, with finite mean, μ_R, and deviation, σ_R, let R^* be the random variable

$$R^* ::= \frac{R - \mu_R}{\sigma_R}.$$

R^* is called the "normalized" version of R.

Definition. The *normal density function* is the function

$$\eta(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

and the *normal distribution function* is its integral

$$N(y) = \int_{-\infty}^{y} \eta(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} dx.$$

The function $\eta(x)$ defines the standard *Bell curve*, centered about the origin with height $1/\sqrt{2\pi}$ and about two-thirds of its area within unit distance of the origin. The normal distribution function $N(y)$ approaches 0 as $y \to -\infty$. As y approaches zero from below, $N(y)$ grows rapidly towards $1/2$. Then as y continues to increase beyond zero, $N(y)$ rapidly approaches 1.

Theorem (Central Limit Theorem). *Let $S_n = \sum_{i=1}^{n} X_i$ where $X_1, \ldots, X_i, \ldots$ are mutually independent variables with the same mean, μ, and deviation, σ, and let S_n^* be the normalized version of S_n. Then*

$$\lim_{n \to \infty} \Pr\{S_n^* \leq \beta\} = N(\beta) \tag{4}$$

for any real number β.