Math/Stat 251 Fall 2015
The Law of Large Numbers

The law of large numbers, sometimes called the law of averages, is precisely the result that casinos rely on in order to make money! The basic idea is that the *sample average is a very good approximation to the population average*. Formally, suppose that $X_1, X_2, \ldots, X_n$ are iid with common mean μ and common variance σ^2. We know from the central limit theorem that if

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

(which is often called the *sample average*), then the limiting distribution of

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

is $\mathcal{N}(0,1)$. The way the central limit theorem is often used in practice is by saying that

$$\overline{X} \text{ is approximately } \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

for large n. It is important to stress that this is not a careful mathematical statement like the central limit theorem. However, it can be shown to provide a useful heuristic. Suppose that we let $n \to \infty$. What happens to the variance of $\overline{X}$? Since

$$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$$

we see that $\mathrm{Var}(\overline{X}) \to 0$ as $n \to \infty$. This suggests that

$$\lim_{n\to\infty} \overline{X} = \lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} \sim \mathcal{N}(\mu, 0).$$

But what does it mean for a normal random variable to have variance 0? It means that there is no *spread* to the distribution. In other words, the density function is concentrated on the single point μ. Thus, this suggests that

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu.$$

The law of large numbers makes this precise.

Theorem (Weak Law of Large Numbers)**.** *Suppose that $X_1, X_2, \ldots, X_n$ are iid with common mean μ and common variance σ^2. Let*

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

so that $\mathbb{E}(\overline{X}) = \mu$ and $\mathrm{Var}(\overline{X}) = \sigma^2/n$. If $\epsilon > 0$ is any positive number, then

$$\lim_{n\to\infty} \mathbf{P}\left(|\overline{X} - \mu| \geq \epsilon\right) = 0.$$

Proof. The idea is to use Chebychev's inequality, namely

$$\mathbf{P}\left(|\overline{X} - \mu| \geq \epsilon\right) \leq \frac{\operatorname{Var}(\overline{X} - \mu)}{\epsilon^2} = \frac{\operatorname{Var}(\overline{X})}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

Hence, taking a limit as $n \to \infty$ gives

$$\lim_{n\to\infty} \mathbf{P}\left(|\overline{X} - \mu| \geq \epsilon\right) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0.$$

Since a probability is always non-negative, we can use the squeeze (or sandwich) theorem to conclude

$$\lim_{n\to\infty} \mathbf{P}\left(|\overline{X} - \mu| \geq \epsilon\right) = 0$$

as required. □

The way that a casino gets to use the LLN is as follows. Suppose that X represents the net winnings of an individual player on one particular play of a certain game. If $\mathbb{E}(X) = \mu$ with $\mu < 0$, then the casino expects to make \$$\mu$ per play. Hence, if the game is played n times, then the casino expects to make \$$n\mu$. Moreover, if $\operatorname{Var}(X) = \sigma^2$, then $\operatorname{Var}(\overline{X}) = \sigma^2/n$. Thus, for n large, there is essentially no variability to $\overline{X}$. The probability that the casino does not make roughly \$$n\mu$ on n plays is therefore negligible. (Chebychev's inequality can be used to give the precise bounds.)

Example. Suppose that a player in a casino is making bets on red at roulette. Hence, for each \$1 bet, the player wins an additional \$1 with probability 18/38 and loses that \$1 with probability 20/38. Let X denote the player's net winnings so that

$$\mathbf{P}(X = 1) = \frac{18}{38} \quad \text{and} \quad \mathbf{P}(X = -1) = \frac{20}{38}.$$

Furthermore,

$$\mathbb{E}(X) = (1) \cdot \frac{18}{38} + (-1) \cdot \frac{20}{38} = -\frac{2}{38} = -\frac{1}{19}.$$

Now, suppose that $X_1, X_2, \ldots$ are iid each with this common distribution so that their common mean is $\mu = -1/19$. In other words, we can think of X_j as a \$1 bet by player j. The law of large numbers tells us that

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = -\frac{1}{19}.$$

To put this another way,

$$X_1 + X_2 + \cdots + X_n \approx -\frac{n}{19}.$$

That is, if n bets are made this way at roulette, then the casino expects to make \$$n/19$. Since

$$\mathbb{E}(X^2) = (1)^2 \cdot \frac{18}{38} + (-1)^2 \cdot \frac{20}{38} = 1,$$

we conclude

$$\sigma^2 = \mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 1 - \left(-\frac{1}{19}\right)^2 = \frac{360}{361}.$$

Therefore, by Chebychev's inequality,

$$\mathbf{P}\left(|\overline{X} - \mu| \geq \epsilon\right) = \mathbf{P}\left(\left|\overline{X} + \frac{1}{19}\right| \geq \epsilon\right) \leq \frac{360}{361 n \epsilon^2},$$

or equivalently,

$$\mathbf{P}\left(|(X_1 + \cdots + X_n) + \frac{n}{19}| \geq n\epsilon\right) \leq \frac{360}{361 n \epsilon^2}.$$

Suppose that $n = 1,000,000$ (not at all unreasonable for a casino) and let $\epsilon = 0.01$ so that

$$\mathbf{P}\left(|(X_1 + \cdots + X_n) + 52631.58| \geq 10000\right) \leq \frac{360}{36100} = 0.0099723.$$

In other words, the probability that the casino's profit will *not* be in the range \$42631.58 and \$62631.58 is 1%; i.e., there is a 99% chance that the casino will make at least \$42000!

Example. Suppose that a player in a casino is making *working 6/8 place bets at craps.* (See Problem #4 on Assignment #6.) This means that for each \$12 bet, the player wins an additional \$7 with probability 5/8 and loses that \$12 with probability 3/8. Hence, if X denotes the player's net winnings, then

$$\mathbf{P}(X = 7) = \frac{5}{8} \quad \text{and} \quad \mathbf{P}(X = -12) = \frac{3}{8}$$

so that

$$\mathbb{E}(X) = (7) \cdot \frac{5}{8} + (-12) \cdot \frac{3}{8} = -\frac{1}{8}.$$

The law of large numbers tells us that if n bets are made this way, then the casino expects to make $\$n/8$.

Remark. This is not a contradiction, but notice in this last example that an individual player is more likely to win money than lose money. This means that more people who make this bet are likely to go away winners than losers. This does not contradict the LLN since the LLN only applies to the casino and not to an individual player. In fact, from a marketing point of view, this is very good for business for the casino. Most people will go away happy since they've won: there will be a lot of small winners, but only a few large losers.

Summary. You can "win" at a casino. However, there are two caveats. You (i) must be willing to tolerate a loss of \$12 in order to have a 5/8 chance of winning \$7, and (ii) must only ever in your lifetime make a single bet; as soon as you start making multiple bets, the LLN applies to you too!