4.3 The Central Limit Theorem

We return to dice rolling for the moment to motivate the next result. Suppose you rolled a die 50 times and recorded the average roll as $\bar{X}_1 = \frac{1}{50}\sum_{k=1}^{50} X_k$. Now you repeat this experiment by rolling the die 50 more times and recording the new average roll as $\bar{X}_2$. You continue doing this and obtain a sequence of sample means $\{\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots\}$. If you plotted a histogram of the results, you would begin to notice that the $\bar{X}_i$'s begin to look normally distributed. What are the mean and variance of this approximate normal distribution? They should agree with the mean and variance of the random variable $\bar{X}_i$, which we compute below. Note that these calculations don't depend on the index i, since each $\bar{X}_i$ is a sample mean computed from 50 independent fair die rolls (all $\bar{X}_i$'s come from the same distribution). As a result, and for ease of notation, we omit the index i and just denote the sample mean as $\bar{X} = \frac{1}{50}\sum_{k=1}^{50} X_k$.

$$
\begin{aligned}
E(\bar{X}) &= E\Big(\frac{1}{50}\sum_{k=1}^{50} X_k\Big) \\
&= \frac{1}{50}\sum_{k=1}^{50} E(X_k) \\
&= \frac{1}{50}\sum_{k=1}^{50} 3.5 \\
&= \frac{1}{50}\cdot 50 \cdot 3.5 \\
&= 3.5
\end{aligned}
$$

where the second equality follows from linearity of expectations, and the third equality follows from the fact that the expected value of a die roll is 3.5 (See Section 2.2). The variance of $\bar{X}_i$ is

$$
\begin{aligned}
\mathrm{Var}(\bar{X}) &= \mathrm{Var}\Big(\frac{1}{50}\sum_{k=1}^{50} X_k\Big) && \text{(Definition of } \bar{X}_i\text{)} \\
&= \frac{1}{50^2}\mathrm{Var}\Big(\sum_{k=1}^{50} X_k\Big) && (\mathrm{Var}(cY) = c^2\mathrm{Var}(Y)) \\
&= \frac{1}{50^2}\sum_{k=1}^{50}\mathrm{Var}(X_k) && (X_k\text{'s are independent.}) \\
&= \frac{1}{50^2}\cdot 50\cdot \mathrm{Var}(X_k) && (X_k\text{'s are identically distributed.}) \\
&\approx \frac{1}{50}\cdot 2.92 \\
&\approx 0.0583
\end{aligned}
$$

where we computed $\mathrm{Var}(X_k) \approx 2.92$ in Exercise 2.12. So we would begin to observe that the sequence of sample means begins to resemble a normal distribution with mean $\mu = 3.5$ and variance $\sigma^2 = 0.0582$. This amazing result follows from the Central Limit Theorem, which is stated below.

Theorem 4.16 (Central Limit Theorem)**.** *Let $X_1, X_2, X_3, \ldots$ be iid (independent and identically distributed) with mean μ and variance $\sigma^2 > 0$. Then*

$$\bar{X} \to N\Big(\mu, \frac{\sigma^2}{n}\Big)$$

in distribution as $n \to \infty$.

All this theorem is saying is that as the number of samples n grows large, independent observations of the sample mean $\bar{X}$ look as though they were drawn from a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$. The beauty of this result is that this type of convergence to the normal distribution holds for any underlying distribution of the X_i's. In the previous discussion, we assumed that each X_i was a die roll, so that the underlying distribution was discrete uniform over the set $\Omega = \{1, 2, 3, 4, 5, 6\}$. However, the convergence result of the CLT is true for just about any underlying distribution of the X_i's. The only requirements are that the X_i's must have positive variance (they can't simply be constants) and the mean and variance must both be finite (this is usually the case).

A continuous distribution we have not yet discussed is the Beta distribution. It is characterized by two parameters α and β (much like the normal distribution is characterized by the parameters μ and σ^2). On the Central Limit Theorem page of the website, choose values for α and β and observe that the sample means look as though they are normally distributed. This may take a while but continue sampling (hitting "Submit") until the histogram begins to fit the normal curve (click the check box next to "Theoretical" to show the plot of the normal curve).

Corollary 4.17. *Another way to write the convergence result of the Central Limit Theorem is*

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \to N(0, 1)$$

Proof. By the CLT, $\bar{X}$ becomes distributed $N(\mu, \frac{\sigma^2}{n})$. By Proposition 4.15 (c), $\bar{X} - \mu$ is then distributed

$$\bar{X} - \mu \sim N\Big(\mu - \mu, \frac{\sigma^2}{n}\Big) = N\Big(0, \frac{\sigma^2}{n}\Big)$$

Combining the above with Proposition 4.15 (a), we have that $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ is distributed

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N\Big(0, \frac{\sigma^2}{n} \cdot \Big(\frac{1}{\sigma/\sqrt{n}}\Big)^2\Big) = N(0, 1)$$

$\square$