5.6. The Central Limit Theorem

1. *Theorem.* Let $X_1, \cdots, X_n$ be a random sample of size n from a population X, and suppose that the population X has mean μ and variance σ^2. Define the sample mean $\bar{X}$ as usual, that is,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\left(X_1 + X_2 + \cdots + X_n\right).$$

   It is clear that $\bar{X}$ has mean μ and variance σ^2/n. Let W be the standardization of $\bar{X}$. That is, define

$$W = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

   Then W has mean 0 and variance 1. If n is large, then the distribution of W is approximately $N(0, 1)$.

* The proof is out of the scope of this course and is omitted.

2. *Example.* Suppose that $X_1, X_2, \cdots, X_{90}$ are independent $Exponential(1)$ distributions. Define the sample sum Y as usual, that is,

$$Y = X_1 + X_2 + \cdots + X_{90}.$$

Then, Y has mean 90 and variance 90. Define W as the standardization of Y. That is,

$$W = \frac{Y - 90}{\sqrt{90}}.$$

Then, $W \approx N(0,1)$.

* In fact, since $Y = X_1 + X_2 + \cdots + X_{90}$ and each $X_i \sim Exponential(1)$, we have

$$Y \sim Gamma(90, 1).$$

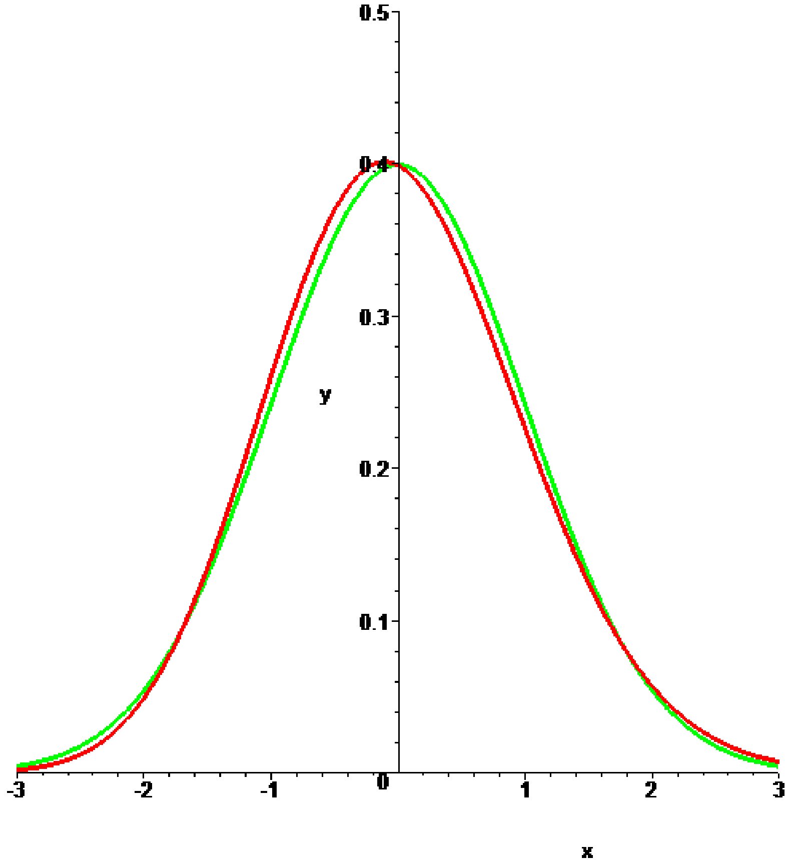
Hence, again, $E(Y) = 90$, $Var(Y) = 90$. The standardization of Y is

$$W = \frac{Y - 90}{\sqrt{90}}.$$

Since 90 is (somewhat) large, by the central limit theorem,

$$W = \frac{Y - 90}{\sqrt{90}} \approx N(0, 1).$$

* The pdfs of $W = \frac{Y-90}{\sqrt{90}}$ and $N(0, 1)$ are plotted on the next page. You will see that they are indeed close to each other.


0.5
0.4
0.3
y
0.2
0.1
-3
-2
-1
0
1
2
3
x