Theorem (Central Limit Theorem)**.** If $X_1, X_2, \ldots, X_n$ constitute a random sample from a population with mean μ and variance σ^2, then the limiting distribution of

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

as $n \to \infty$ is the standard normal distribution.

Lemma. If a and b are constants, then $M_{aX+b}(t) = e^{bt} M_X(at)$.

1. Suppose that Z in the CLT is actually a standard normal random variable. Starting with the moment-generating function of Z, find the moment-generating functions of $\overline{X}$ and of $T = n\overline{X}$. Identify the distributions of these two random variables (these will be approximations, since Z is only approximately standard normal).

2. Of the 270 numbers you collected last week, 92 started with the digit 1. I think that each digit should be equally likely to be the first digit, so the digit 1 should be the first digit exactly $\frac{1}{9}$ of the time. Assume that our numbers represent a random sample.

a) Estimate the probability of 92 or more of the first digits being 1 using a normal approximation.

b) Calculate the probability of 92 or more of the first digits being a 1 using a binomial cumulative distribution function (on a computer or calculator).

3. Of the 258 second digits you collected, 20 had a second digit of 1. Repeat the analysis of the previous problem for second digits (which I think should be 1 exactly $\frac{1}{10}$ of the time). Does it matter if you use a correction for continuity?

Lemma. If $X_1, X_2, \ldots, X_n$ are independent random variables and X_i has moment-generating function $M_{X_i}(t)$, then

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t)$$

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a population having a Poisson distribution with parameter λ.

a) Use moment-generating functions to identify the distribution of $T = \sum_{i=1}^n X_i$.

b) Combine part a with problem 1 to argue that when λ is large, the Poisson distribution may be approximated by a normal distribution.

c) Explore the accuracy of your approximation by comparing exact values and estimates for $P(T \leq \lambda)$ when $\lambda = 5, 10, 15, 20, 25, 30$. (Would a correction for continuity be appropriate?)