: 5.3

F. Central Limit Theorem:

1. If a large collection of n random variables X_j all have the same distribution and are independent independent and identically distributed (iid), then their mean is approximately normally distributed, if Var $[X_j] < \infty$.
2. Since the only difference between the mean and the sum is division by n, then the same holds for the sum
3. Pedantic version:
   a. Suppose $X_1, \ldots, X_n$ iid.
   b. Let $Z = \Sigma_{j=1}^{n} X_j - n\mu)/(\sqrt{n}\sigma)$
   c. Then $\lim_{n\to\infty} P(Z \le z) \to \Phi(z)$.
4. Example: 100 subjects are given two medications in turn (ex. two pain relievers) and are asked which did a better job for them. 57 of them chose medication A as better. Is it plausible that both medications are equally effective?
5. $X \sim$ Bin$(100, .5)$; $P(X \geq 57) \approx 1 - \Phi((57 - 50)/\sqrt{100 \times .5 \times .5}) = 1 - \Phi(7/5) = 1 - .92 = .08$.
6. Why?

a. Suppose $\mu = 0$, $\sigma = 1$.

b. Theorem says that as n increases, the distribution of Z should get closer to symmetric.

c. Let's use $\mathrm{E}[Z^3]$ (called kurtosis) as measure of *skewness* (asymmetry).

glossa
skewn

d. $Z^3 = \Sigma_{j=1}^{n} X_j^3/(n\sqrt{n}) +$ terms that involve either one of the X_j s squared and another X_k, or a product of three different X_j s.

e. Expectation of product of independent X_j s is product of expectations.

f. So cross terms have zero expectation.

g. So $\mathrm{E}[Z^3] = \Sigma_{j=1}^{n} \mathrm{E}[X_j]/(n\sqrt{n}) = \Sigma_{j=1}^{n} \mathrm{E}[X_1]/(n\sqrt{n}) = \mathrm{E}[X_1]/\sqrt{n} \to 0$.

h. Similar arguments hold for other measures of differences between the distribution of Z and $\mathsf{N}(0,1)$.