STAT 312 Lab 10

1. (a) State the Central Limit Theorem.

   (b) Let $S_n \sim bin(n,p)$. By thinking of S_n as the number of 'successes', represent it as a sum of independent r.v.s (recall that we did exactly this when we discussed the binomial p.g.f. and m.g.f.) and conclude that

$$\frac{S_n - np}{\sqrt{n}} \xrightarrow{L} N\left(0, p\left(1-p\right)\right).$$

   (a) Recall the 'delta method'; state the approximate mean and variance of a function $\psi\left(\bar{Y}_n\right)$ of a sample average $\bar{Y}_n = n^{-1}\sum Y_i$, when the Y_i each have mean μ_Y and variance σ_Y^2

   (b) Suppose that $\hat{p}$ is the proportion of n individuals exhibiting a certain trait, where p is the probability that an individual exhibits this trait, and let $Z_n = \log\left(\hat{p}/(1-\hat{p})\right)$ be the 'logit'. This is sometimes preferred to $\hat{p}$ since it is not constrained to lie in $(0,1)$ – it can take on any value and in fact is approximately normally distributed, as you will now show. More precisely, show that

$$\sqrt{n}\left(Z_n - \log\left(\frac{p}{1-p}\right)\right) \xrightarrow{L} N\left(0, \frac{1}{p\left(1-p\right)}\right).$$

2. Suppose we gather a sample $X_1, ..., X_n$ of i.i.d. observations, obtaining the numerical values $x_1, ..., x_n$.

   (a) If the X_i have the d.f. F, then a common estimate of F is the *empirical distribution function* (e.d.f.)

$$\hat{F}_n(x) = \frac{\#\text{ of } X_i \text{ which are } \leq x}{n}.$$

   Make a plot of $\hat{F}_n(x)$. Show that $\hat{F}_n(x)$ is the average of the i.i.d. r.v.s $Z_i = I\left(X_i \leq x\right)$, each of which has a Bernoulli distribution with $p = F(x)$.

   (b) Apply the WLLN to assert that, for each x, $\hat{F}_n(x) \xrightarrow{pr} F(x)$.

   (c) Apply the CLT so as to exhibit an appropriately normalized version of $\hat{F}_n(x)$ which has a limiting Normal distribution.

Here we have looked at $\hat{F}_n(x)$ for fixed x; as x varies we obtain the *empirical process*, which is one of the most well-studied stochastic processes in Probability Theory.

...over

3. (a) State the Weak Law of Large Numbers.

   (b) Let $X_1, ..., X_n$ be a sample from a population with mean μ and variance σ^2. Complete the 'application' of Lecture 29 by showing that $S^2 \xrightarrow{pr} \sigma^2$. [*Hint*: start by showing that

$$S^2 = \frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - \left(\bar{X}-\mu\right)^2\right];$$

   then think about the WLLN and Slutsky's Theorem.]

4. Recall, from Lecture 24, the m.g.f. of a negative binomial r.v. Suppose that $p = p_r$ increases with r, in such a way that

$$r\frac{1-p_r}{p_r} \to \lambda > 0 \text{ as } r \to \infty.$$

Thus 'successes' are increasingly likely, but more of them are required before the trials can stop. Here you will show that the number of failures before the r^{th} success, which we now write as N_r, has a limiting $\mathbb{P}(\lambda)$ distribution.

   (a) First show that the m.g.f. can be written as

$$E\left[e^{tN_r}\right] = \exp\left\{-r\log\left(1-c_r\right)\right\},$$

   where $c_r = \frac{1-p_r}{p_r}\left(e^t-1\right)$. Note that $c_r \to 0$ as $r \to \infty$.

   (b) Argue that $N_r \xrightarrow{L} \mathbb{P}(\lambda)$ iff $-r\log(1-c_r) \to \lambda\left(e^t-1\right)$.

   (c) Establish the requirement of (b).

5. Here you will show that the standard normal density $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ really is a density, in that it integrates to 1; i.e. you will show that if we define $I = \int_{-\infty}^{\infty} e^{-x^2/2}dx$, then $I = \sqrt{2\pi}$. For this, write I^2 as

$$I^2 = \int_{-\infty}^{\infty} e^{-x_1^2/2}dx_1 \cdot \int_{-\infty}^{\infty} e^{-x_2^2/2}dx_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\left(x_1^2+x_2^2\right)/2}dx_1dx_2,$$

and transform to polar coordinates as described in Lecture 30.