Sta 250 = Mth 342 : Homework 5

1. A random sample $\mathbf{x} = \{X_i : 1 \le i \le n\}$ is taken from the uniform distribution $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{Un}(0,\theta)$ on the interval $[0,\theta]$ for some unknown $\theta > 0$. Let $\delta(\mathbf{x}) = \max\{X_i\}$ be the maximum observation. How large must n be to ensure that

$$\mathsf{P}\left[\ |\delta(\mathbf{x}) - \theta| < \frac{\theta}{10}\ \right] \ge 0.90$$

for every $\theta > 0$?

2. A random sample $\{X_i : 1 \le i \le n\}$ is taken from the Poisson distribution $X_i \overset{\text{iid}}{\sim} \mathsf{Po}(\theta)$ with unknown mean $\theta > 0$. How large must n must be to ensure that

$$\mathsf{E}_\theta[\ |\bar{X}_n - \theta|^2/\theta\ ] \le \frac{1}{10}$$

for every $\theta > 0$?

3. If the sample size n in Problem 2 is large enough, the central limit theorem will ensure that $\sum_{i \le n} X_i$ and hence $\bar{X}_n = (\sum_{i \le n} X_i)/n$ will have approximately normal distributions. How large must n must be to ensure that

$$\mathsf{P}\left[\ |\bar{X}_n - \theta| < \frac{1}{10}\sqrt{\theta}\ \right] \ge 0.90$$

for every $\theta > 0$, assuming the normal approximation holds?

4. Still with the same sample as in Problems 2 and 3, how large must n must be to ensure that

$$\mathsf{P}_\theta\left[\ |\bar{X}_n - \theta| < 0.01\ \right] \ge 0.90$$

for the specific value of $\theta = 4$?

5. The observations $\mathbf{x} = \{X_1, \ldots, X_n\}$ of stearic acid concentration in clams may be treated as independent normal random variables $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{No}(\mu, \sigma^2)$ with unknown mean and variance. If the mean μ is known then one estimate of the variance σ^2 is

$$S = \frac{1}{n}\sum_{i \le n}(X_i - \mu)^2$$

If $n = 8$ and $\sigma = 0.5$, what number c satisfies

$$\mathsf{P}(S \le c) = 0.90?$$

6. Suppose in Problem 5 we had $n = 200$ observations. Again with $\sigma = 0.5$, find (approximately, using the central limit theorem) the smallest number c that satisfies

$$\mathsf{P}(S \le c) = 0.90$$

Justify your use of the central limit theorem.

7. For $n = 1$, $n = 2$ and $n = 4$, *sketch* a plot of the $\chi^2_n = \mathsf{Ga}(n/2, 1/2)$ distribution with n degrees of freedom. Find and indicate on your plots the *mean*, *mode*, and *median* of each.

8. Let $X, Y \overset{\text{iid}}{\sim} \mathsf{No}(0, 1)$ be independent random variables, each with the standard normal distribution. Find the expected area of the smallest circle centered at the origin that contains the point (X, Y).

9. In 1827 botanist Robert Brown noticed that pollen grains seemed to wiggle when observed under a microscope— some of the first direct evidence that molecules exist. Each of the three coordinates of the position of a pollen grain may be viewed as the sum of thousands and thousands of "kicks" the grain gets from its collisions with water molecules. By the central limit theorem, each coordinate has approximately a normal distribution, all independent. Call the coordinates at time $t \geq 0$ X_t, Y_t, and Z_t, and fix the origin of the coordinate system so that $(X_0, Y_0, Z_0) = (0, 0, 0)$. Then for some constant σ, after any number t of seconds of jostling by water molecules, each coordinate will have approximately a $\mathsf{No}(0, \sigma^2 t)$ distribution. Einstein used estimates of σ to find the first estimates of the mass of water molecules (and hence the first estimates of Avagadro's number).

   Find the probability that at time $t = 3$ the pollen grain will lie inside the sphere whose center is the origin and whose radius is 5σ.

10. Let $\hat{\mu}$ and $\hat{\sigma}^2$ be the maximum likelihood estimates of the mean μ and variance σ^2 of a normal distribution, based on a random sample $\mathbf{x} = \{X_1, \ldots, X_n\} \overset{\text{iid}}{\sim} \mathsf{No}(\mu, \sigma^2)$. Show that $\hat{\sigma}^2 \sim \mathsf{Ga}(\alpha, \lambda)$ has a Gamma probability distribution, and find α and λ in terms of μ, σ^2, and n.

Note: Earlier problems 11+ have been removed.