**ECON 2222A Thursday, November 5: warm-up**

We'll start class Thursday by discussing two practice problems I developed based on my grading of Quiz 3. Please attempt to work through each ahead of class, and I'll choose one of you to go up to the board and present your solution (we'll help!). Show your reasoning, not just the numbers; the "why" parts matter most.

**1. Number of successes, and the sample proportion.** In the wage example (Oct 29), the share of workers earning \$50 or more is $\pi = 0.20$. Take a random sample of $n = 400$ workers. Code worker $i$ as $X_i = 1$ if they earn \$50 or more and $X_i = 0$ otherwise, and let $X = \sum X_i$ be the number in the sample earning \$50 or more.

(a) What is the distribution of $X$? Name it (for example normal, Binomial, Bernoulli) and provide its parameters.

(b) Can the distribution of $X$ be approximated by a normal distribution? Show the value you check. If it can, provide the mean, variance and standard deviation of $X$.

(c) The sample mean of the $X_i$ is the share of the sample earning \$50 or more. We will refer to this special version of the sample mean $\overline{X}$ as the *sample proportion*: $p = \frac{X}{n}$ (please note we're using lowercase $p$ to refer to the RV itself and its realization, as this was the minimally confusing notation I could come up with...). What is the approximate distribution of $p$, and why? Provide its mean and standard deviation.

(d) Find $\Pr(X \leq 96)$. (Quiz 3 Q2 asked about this kind of probability.)

(e) Find $\Pr(p \leq 0.24)$. Write both (d) and (e) as a probability about $Z$. Why do they agree?

**2. A difference between two normals.** Your one-way commute is $X \sim N(30, 16)$ and a friend's is $Y \sim N(27, 9)$, in minutes, and is independently distributed from your commute. Let $D = X - Y$.

(a) What is the distribution of $D$, and why? Name the distribution (for example normal, Binomial, uniform), then provide its mean, variance and standard deviation.

(b) Find $\Pr(X > Y)$. Write it first as a probability about $D$, then about $Z$.

(c) A common Quiz 3 answer was $Var(D) = 16 - 9 = 7$. What is wrong with it? (Hint: what if we had defined the difference as $Y - X$ instead?)

(d) Now average over days: $\overline{X}$ is your mean commute over $n_X = 4$ days and $\overline{Y}$ your friend's mean over $n_Y = 9$ days (all days independent). What is the distribution of $\overline{X} - \overline{Y}$, and why? Find $\Pr(\overline{X} > \overline{Y})$.