MATH SUPPLEMENT ON NORMAL PROBABILITY PLOTS

(for pp. 129 – 130 of *Stats: Data and Models*, DeVeaux, Velleman and Bock, 3^{rd} edition.)

1. **Connection with M 362K prerequisite**: The phrase "what value to expect for the smallest z-score in a batch of 100" (p. 130) is talking about the expected value $E(Z_1)$ of the random variable Z_1 defined as:

Z_1 = the smallest z-score in a sample of 100 from a certain random variable.

2. **Why the normal probability plot will be close to a straight line if the underlying distribution is normal:**

Label the data $y_1, y_2, \dots , y_n$ from smallest to largest – so

$$y_1 \leq y_2 \leq \dots \leq y_n$$

Label the normal scores $z_1, z_2, \dots , z_n$ from smallest to largest – so

$$z_1 \leq z_2 \leq \dots \leq z_n$$

If the y's are indeed from a normal distribution, say with mean μ and standard deviation σ , then if we standardize the y's (but using μ and σ instead of the sample mean and sample standard deviation), the resulting z-score formed from y_i should be approximately equal to the i^{th} normal score z_i:

$$(y_i - \mu)/ \sigma \approx z_i$$

Solving for y_i gives:

$$y_i \approx \sigma z_I + \mu$$

In other words, plotting the y_i's (data) against the z_i's (normal scores) to get the normal probability plot should give approximately a straight line with slope = _____ and y-intercept = _____.

Questions to think about:

1. Can you use the above to explain why the normal plot from a distribution that is skewed to the right should have a shape like that in Figure 6.10 on p. 129?
2. What if the distribution is skewed to the left?
3. What might a normal probability plot from a uniform distribution look like?