THE NORMAL AS A FOUNDATION FOR STATISTICAL INFERENCE – USING RELATIONSHIPS AMONG Z-SCORES, PERCENTILE RANK, AND AREA UNDER THE NORMAL CURVE

A. Spatz (1997, pp. 369-370, Table C)

Refer to Table C in Spatz (1997, pp 369-370) throughout

B. Using the table

1. we can use the table and accompanying drawings to determine, for example, the percentile rank (PR) of an observation in a normal distribution if we know its z-score

a. use z-score and then look at Column C in Spatz's Table C

b. recall that we can generate z-scores for observation in **any** kind of distribution – we can use the relationship among z, PR, and area under the normal (as shown in Table C of Spatz) on **raw scores** IFF (if and only if) the original distribution is normal

Using the table (cont'd)

2. Given a normal distribution

What proportion of scores is below an observation with $z = 1.9$? i.e., what is the percentile rank (PR) of an observation 1.9 standard deviations above the mean in any normal distribution?

a. from Table C, $z = 1.9 \Rightarrow 0.4713$ between μ and that z-score, i.e., 47.13% of the area under the curve, and, thus, 47.13% of the values in the normal distribution, lie between μ and that z-score

$0.4713 + 0.5000 = 0.9713$, 97.13%, 97.13^{th} percentile

Using the table (cont'd)

b. 1.0000 – 0.0287 = 97.13% (from Column C)

Refer to Table C in Spatz (1997, pp. 369-370)

Using the table (cont'd)

3. Given a normal distribution

   What is the PR of a score where $z = 0.37$? In other words, what proportion of scores/what proportion of the area under the normal lies below an observation 0.37 standard deviations above the mean?

Table C, Column B, $z = 0.37 \Rightarrow 0.1443$

Using the table (cont'd)

a. 0.1443 + 0.5000 = 0.6443, 64.43%, 64.43rd percentile

1.0000 – 0.3557 = 64.43%

Refer to Table C in Spatz (1997, pp. 369-370)

Using the table (cont'd)

4. what about the PR of a score in a normal distribution where $z = -0.37$?

a. Table C, Column B, $z = -0.37 \Rightarrow 0.1443$

$-0.1443 + 0.5000 = 0.3557$, 35.57%, 35.57th percentile

Using the table (cont'd)

b. Table C, Column C directly – 0.3557

C. How do we determine the PR of a raw score in a normal distribution if we do not know z?

1. an easy process

a. determine z
b. use Table C, Column B, add to or subtract from 0.5000 as appropriate

How do we determine the PR of a raw score (cont'd)

c. use Table C, Column C as appropriate

(i.) if z>0, subtract from 1.0000

(ii) if z<0, just read column C

How do we determine the PR of a raw score (cont'd)

2. we will use the Stanford-Binet IQ test as an example, despite all that we know about the weaknesses of standardized tests especially this one

a. a normal distribution is sometimes identified by this convention

ND (μ, σ), sometimes written ND (μ,σ^2) – recall that every normal distribution is uniquely identified by its mean and standard deviation

ND (100, 15) – what that means is that the distribution of scores on the Stanford-Binet IQ test is normally distributed, with a mean of 100 and a standard deviation equal to 15

How do we determine the PR of a raw score (cont'd)

b. what is the PR of an IQ score of 119?

(i) determine z: $z = \frac{x - \mu}{\sigma} = \frac{119 - 100}{15} = 1.27$

(ii) Table C, Column B, z = 1.27 ⇒ 0.3980

(iii) PR = 0.5000 + 0.3980 = 89.80%

(iv) Using Column C, z = 1.27, area beyond z = 0.1020

(v) PR = 1.0000 – 0.1020 = 0.8980 (Column C)

How do we determine the PR of a raw score (cont'd)

3. ND (22, 1.51), what is the PR of an observation of 18?

OR

What proportion of the area under this normal curve lies below a score of 18?

determine z: $z = \frac{x - \mu}{\sigma} = \frac{18 - 22}{1.51} = 2.65$

How do we determine the PR of a raw score (cont'd)

b. Table C, Column B, $z = -2.65 \Rightarrow -0.4960$

c. $0.5000 - 0.4960 = 0.0040$, 0.40%, 0.40th percentile

d. Table C, column C, $z = -2.65$, area beyond $z = 0.0040$

D. Turn the question around – if we know the percentile rank of a score in a normal distribution, can we determine what that value is?

1. again, a simple process

2. if PR > 50%, subtract 0.5000 from the PR

   use Column B, find the z-score, then use the z-score formula to solve for x

3. if PR < 50%

   use column C, find z-score, then solve z-score formula for x

If we know the percentile rank of a score in a normal distribution (cont'd)

4. ND(100, 15) IQ – what score is at the 89.80th percentile? That is, what value in the dataset is larger than 89.80% of the values in the dataset, or has a PR of 89.80%?

   a. Refer to Spatz, Table C

      0.8980 – 0.5000 = 0. 3980

   b. from Table C, Column B, 0.3980 1.27 = z

If we know the percentile rank of a score in a normal distribution (cont'd)

c. $$z = \frac{x - \mu}{\sigma}$$

$$1.27 = \frac{x - 100}{15}$$

$$(15)(1.27) = x - 100$$

$$19.05 + 100 = x$$

$$x \approx 119$$

remember that scores on the S-B IQ test are in integers (whole numbers)

If we know the percentile rank of a score in a normal distribution (cont'd)

5. ND(10,2) – what observation has a PR of 23%? In other words, what observation is larger than 23% of the values in the data set?

   a. 0.2300 $\Rightarrow$ Column C – no specific value there, but **do not interpolate**

   instead, go to the next highest z-score (farther from the mean)

   b. -0.2296 $\Rightarrow$ -0.74 = z

If we know the percentile rank of a score in a normal distribution (cont'd)

c. $z = \frac{x - \mu}{\sigma}$

d. $-0.74 = \frac{x - 10}{2}$

e. x=8.52

f. generally speaking with regard to normal distributions, $x = \mu + z(\sigma)$