AP® STATISTICS
2008 SCORING GUIDELINES (Form B)

Question 5

Intent of Question

The primary goals of this question were to assess a student's ability to (1) describe the distribution of the difference of two normal random variables and (2) use this distribution to find a probability and to find a value given its location in the distribution.

Solution

Part (a):

X is normally distributed with $\mu = 170$ and $\sigma = 20$, and Y is normally distributed with $\mu = 200$ and $\sigma = 10$. The distribution of $Y - X$ has mean and standard deviation:

$$\mu_{Y-X} = \mu_Y - \mu_X = 200 - 170 = 30$$
$$\sigma_{Y-X} = \sqrt{\sigma_Y^2 + \sigma_X^2} = \sqrt{10^2 + 20^2} = \sqrt{500} = 22.36$$

The distribution of $Y - X$ is normal with mean 30 and standard deviation 22.36 (or, variance 500).

Part (b):

The train from Bullsnake will have to wait when $Y - X$ is positive:

$$P(Y - X > 0) = P\left(z > \frac{0-30}{22.36}\right) = P(z > -1.34) = 0.9099$$

(Calculator: 0.9082408019 or 0.9078172963, if z is not rounded.)
The proportion of days that the train will have to wait is about 0.91.

Question 5 (continued)

Part (c):

Let *D* denote the delay that will be needed for the train leaving Bullsnake. With the additional constant delay,

$X + D$ is normally distributed with $\mu_{X+D} = 170 + D$ and $\sigma_{X+D} = \sigma_X = 20$
Y is normally distributed with $\mu_Y = 200$ and $\sigma_Y = 10$

Thus, the difference $Y - (X + D)$ is normally distributed with

$$\mu_{Y-(X+D)} = \mu_Y - \mu_{(X+D)} = 200 - (D + 170) = 30 - D$$
$$\sigma_{Y-(X+D)} = \sigma_{Y-X} = 22.36$$

The combined delay and travel time ($X + D$) for the Bullsnake train must be *less than* the travel time for the Diamondback train (Y) with probability 0.01. That is, $P(Y - \left(X + D\right) > 0) = 0.01$, so we need

$$\frac{0 - \mu_{Y-(X+D)}}{\sigma_{Y-(X+D)}} = \frac{0 - (30 - D)}{22.36} = 2.33$$

Solving for *D*, the train from Bullsnake should be delayed by 82.099 minutes.

OR, with alternative notation:

Let X' denote the combined delay and travel time from Bullsnake to Copperhead, and let *Y* represent the travel time to Copperhead for the Diamondback train. The distribution of $Y - X'$ also is normal (because *D* is constant), with mean $\mu_{Y-X'} = \mu_Y - \mu_{X'}$ and standard deviation $\sigma_{Y-X'} = \sigma_{Y-X} = 22.36$.

The combined delay and travel time for the Bullsnake train (X') must be *less than* the time for the Diamondback train (Y) with probability 0.01. That is, $P(Y - X' > 0) = 0.01$, and we need

$$z = \frac{0 - \mu_{Y-X'}}{\sigma_{Y-X'}} = \frac{0 - \mu_{Y-X'}}{22.36} = 2.33$$

Solving, $\mu_{Y-X'} = \mu_Y - \mu_{X'} = -52.099$, so the mean travel time for the Diamondback train (Y) should be 52.099 minutes less than the mean combined travel and delay time for the Bullsnake train X'. The mean travel time for the Diamondback train is now 30 minutes more than the mean travel time for the Bullsnake train, so the train from Bullsnake should be delayed by 52.099 + 30 = 82.099 minutes.

Question 5 (continued)

Scoring

Part (a) is divided into two sections: section 1 and section 2. Section 1 is scored as essentially correct (E) or incorrect (I). Section 2 is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response states that the distribution of $Y - X$ is normal.

Incorrect (I) otherwise.

Section 2 is scored as follows:

Essentially correct (E) if the response correctly computes the mean and standard deviation *AND* shows some work for the calculation of the standard deviation. May contain a minor arithmetic error.

Partially correct (P) if the response correctly states the values of the mean and standard deviation.

Incorrect (I) if the formula for the mean or standard deviation contains a conceptual error (such as not squaring the original standard deviations or subtracting the variances).

Part (b) constitutes section 3 and is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 3 is scored as follows:

Essentially correct (E) if the response uses the distribution information from part (a) to correctly compute the desired probability. If the mean or standard deviation is computed incorrectly in part (a), those values should be used in part (b). (*Note:* If variances are incorrectly subtracted instead of added, $\sigma = 17.32$, $z = -1.73$, and the probability is $1 - 0.0418 = 0.9582$.)

Partially correct (P) if the response computes $P(Y - X < 0)$ instead of $P(Y - X > 0)$ and gets $P(Y - X < 0) = 1.0 - 0.9099 = 0.0901$.

Incorrect (I) otherwise.

Part (c) constitutes section 4 and it is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 4 is scored as follows:

Essentially correct (E) if the response correctly concludes that the train from Bullsnake should be delayed by about 82 minutes. If the mean or standard deviation is computed incorrectly in part (a), those values should be used in part (c). (*Note:* If variances are subtracted instead of added, the delay time will be (2.33)(17.32) + 30 = 70.36.)

Partially correct (P) if a correct line of reasoning is explored but the student fails to reach the correct answer.

Incorrect (I) otherwise.

Question 5 (continued)

Each essentially correct response is worth 1 point; each partially correct response is worth ½ point.

4 Complete Response

3 Substantial Response

2 Developing Response

1 Minimal Response

If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down, depending on the strength of the response and communication.