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,

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

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, $P(Y - (X + D) > 0) = 0.01$, 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.

<u>*OR*, with alternative notation</u>:

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**

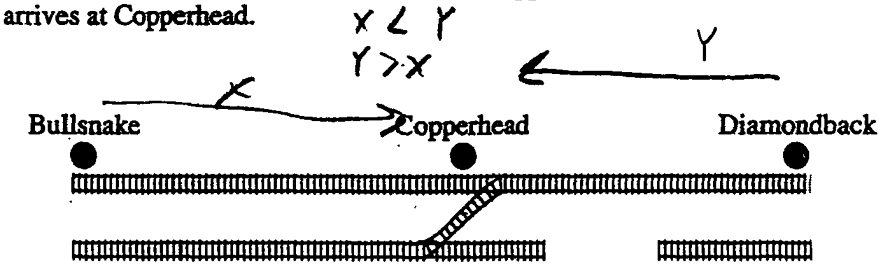
3 **Substantial Response**

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.

5. Flooding has washed out one of the tracks of the Snake Gulch Railroad. The railroad has two parallel tracks from Bullsnake to Copperhead, but only one usable track from Copperhead to Diamondback, as shown in the figure below. Having only one usable track disrupts the usual schedule. Until it is repaired, the washed-out track will remain unusable. If the train leaving Bullsnake arrives at Copperhead first, it has to wait until the train leaving Diamondback arrives at Copperhead.



Every day at noon a train leaves Bullsnake heading for Diamondback and another leaves Diamondback heading for Bullsnake.

Assume that the length of time, X, it takes the train leaving Bullsnake to get to Copperhead is normally distributed with a mean of 170 minutes and a standard deviation of 20 minutes.

Assume that the length of time, Y, it takes the train leaving Diamondback to get to Copperhead is normally distributed with a mean of 200 minutes and a standard deviation of 10 minutes.

These two travel times are independent.

(a) What is the distribution of $Y - X$?

$(Y-X)$ is normally distributed with a mean of 30 minutes and a standard deviation of 22.36 minutes.

$\mu_Y - \mu_X = 200 - 170 = 30$

$\sigma_{Y-X} = \sqrt{\sigma_Y^2 + \sigma_X^2} = \sqrt{20^2+10^2} = 22.36$

Note that both X and Y are independent and normally distributed.

GO ON TO THE NEXT PAGE.

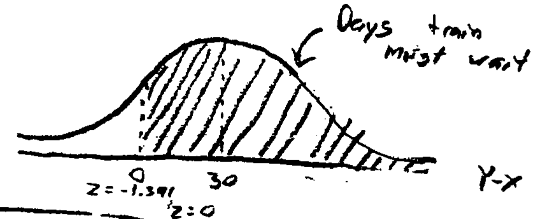
5A2

(b) Over the long run, what proportion of the days will the train from Bullsnake have to wait at Copperhead for the train from Diamondback to arrive?

The train from Bullsnake must wait at Copperhead on these days that Y is greater than X. In other words, it is a day when $Y - X > 0$.

$Y-X$ is normally distributed. $\mu_{Y-X} = 30$ $\sigma_{Y-X} = 22.36$

$z = \frac{x - \mu}{\sigma} = \frac{0-30}{22.36} = -1.341$



$P(z > -1.341) = 1 - .0901 = \boxed{.9099}$

(c) How long should the Snake Gulch Railroad delay the departure of the train from Bullsnake so that the probability that it has to wait is only 0.01 ?

when $P(z) = 1-.01$, $z = 2.33$

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

$2.33 = \frac{0-\mu}{22.36}$

$\mu_{Y-X} = -52.0988$

So, $\mu_Y - \mu_X = -52.0988$

$200 - \mu_X = -52.0988$

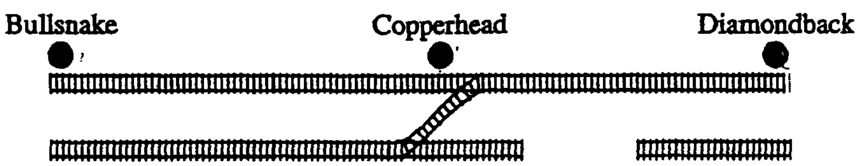
$\mu_X = 252.0988$

Since the travel time of X is 170 minutes, the delay should be $252.0988 - 170 = \boxed{82.0988 \text{ minutes.}}$

Note that this is still a normal distribution.

GO ON TO THE NEXT PAGE.

5. Flooding has washed out one of the tracks of the Snake Gulch Railroad. The railroad has two parallel tracks from Bullsnake to Copperhead, but only one usable track from Copperhead to Diamondback, as shown in the figure below. Having only one usable track disrupts the usual schedule. Until it is repaired, the washed-out track will remain unusable. If the train leaving Bullsnake arrives at Copperhead first, it has to wait until the train leaving Diamondback arrives at Copperhead.



Every day at noon a train leaves Bullsnake heading for Diamondback and another leaves Diamondback heading for Bullsnake.

Assume that the length of time, X, it takes the train leaving Bullsnake to get to Copperhead is normally distributed with a mean of 170 minutes and a standard deviation of 20 minutes.

Assume that the length of time, Y, it takes the train leaving Diamondback to get to Copperhead is normally distributed with a mean of 200 minutes and a standard deviation of 10 minutes.

These two travel times are independent.

(a) What is the distribution of $Y - X$?

$$\mu_{Y-X} = 200 - 170 = 30$$

$$\sigma_{Y-X} = \sqrt{(20)^2 + (10)^2} = 22.36$$

Normally distributed with a mean of 30 minutes and a standard deviation of 22.36 minutes.

GO ON TO THE NEXT PAGE.

5B2

(b) Over the long run, what proportion of the days will the train from Bullsnake have to wait at Copperhead for the train from Diamondback to arrive?

$P(Y-X) > 0$

normalcdf $(0, 10000, 30, 22.36) = .91$

(c) How long should the Snake Gulch Railroad delay the departure of the train from Bullsnake so that the probability that it has to wait is only 0.01 ?

normalcdf $(0, 10,000, -50, 22.36) = .01$

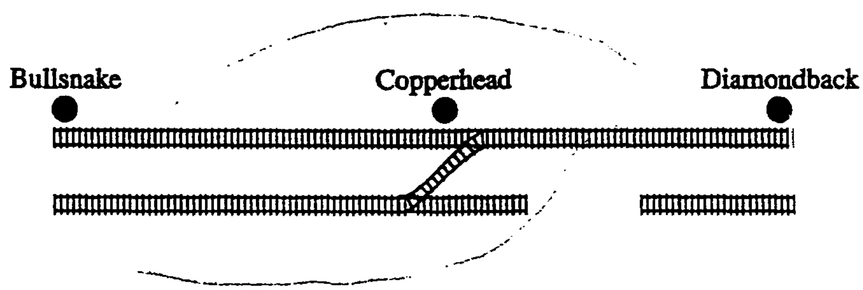
$200 - x = -50$

$x = 250$

wait at least 80 minutes

GO ON TO THE NEXT PAGE.

5. Flooding has washed out one of the tracks of the Snake Gulch Railroad. The railroad has two parallel tracks from Bullsnake to Copperhead, but only one usable track from Copperhead to Diamondback, as shown in the figure below. Having only one usable track disrupts the usual schedule. Until it is repaired, the washed-out track will remain unusable. If the train leaving Bullsnake arrives at Copperhead first, it has to wait until the train leaving Diamondback arrives at Copperhead.



Every day at noon a train leaves Bullsnake heading for Diamondback and another leaves Diamondback heading for Bullsnake.

Assume that the length of time, X, it takes the train leaving Bullsnake to get to Copperhead is normally distributed with a mean of 170 minutes and a standard deviation of 20 minutes.

Assume that the length of time, Y, it takes the train leaving Diamondback to get to Copperhead is normally distributed with a mean of 200 minutes and a standard deviation of 10 minutes.

These two travel times are independent.

(a) What is the distribution of $Y - X$?

Since they are independent:

mean $(Y-X) = 200 - 170 = 30$ minutes

variance $(Y+X) = 100 + 400 = 500$

standard deviation $= \sqrt{500} = 22.36$ minutes

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

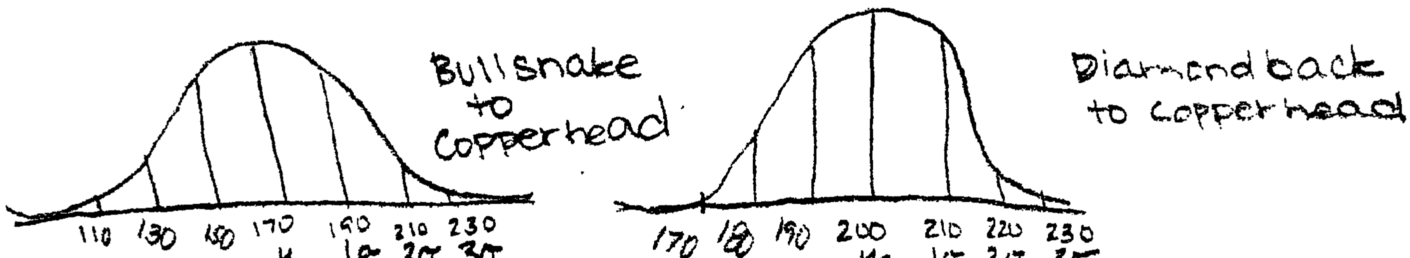
5C2

(b) Over the long run, what proportion of the days will the train from Bullsnake have to wait at Copperhead for the train from Diamondback to arrive?

$$z = \frac{30-0}{22.36} = 1.34 \qquad z = \frac{\bar{x}-\mu_0}{\sigma}$$

$$P = .9099 = \frac{9099}{1000}$$

About 90.99% of the days the train from Bullsnake to Copperhead will have to wait for the train from Diamondback to Copperhead.



(c) How long should the Snake Gulch Railroad delay the departure of the train from Bullsnake so that the probability that it has to wait is only 0.01 ?

z value at .99 = 2.33

$$2.33 = \frac{X-170}{20}$$

$$46.6 = X - 170$$

$$216.6 = X \qquad 216.6 - 170 = 46.6$$

It should delay 46.6 or 47 minutes so that the probability of waiting is .01.

Unauthorized copying or reuse of any part of this page is illegal.

GO ON TO THE NEXT PAGE.

Question 5

Sample: 5A
Score: 4

For almost every section of this question, this essay provides correct conclusions and computations, with justification. The exception is that a clear justification is not given for the statement that "(Y – X) is normally distributed" in part (a). The notation in part (c) could be improved, but the reasoning is correct. Overall, this answer was considered a complete response, based on all four sections that constitute the three parts.

Sample: 5B
Score: 3

Part (a) is done well, with the student providing justification for the mean and standard deviation (but not for normality). Both sections in part (a) were scored as essentially correct. In parts (b) and (c) much of the justification for the answer is based on a calculator command. Explanation for an answer should not depend on the reader's familiarity with a specific calculator. Further, the answer of "at least 80 minutes" rather than 82 minutes in part (c) apparently is an estimate found by using the given calculator command, but it is impossible to know from the information given. Thus, the essay does not indicate clearly the methods used. Section 3 was scored as essentially correct, but section 4 was scored as incorrect. On the whole, this answer was judged a substantial response.

Sample: 5C
Score: 2

The mean and standard deviation of the distribution of $Y - X$ are correctly computed in part (a), but the shape is not described. Section 1 was scored as incorrect, and section 2 was scored as essentially correct. Part (b) (section 3) was scored as essentially correct, while part (c) (section 4) was scored as incorrect. Also, this response did not include a sketch of the distribution of the difference of times in parts (b) or (c). Diagrams are not necessary, but an absence of diagrams is often paired with incorrect computations. Overall, this answer was assessed as a developing response.