MAYO

2003 AP® STATISTICS FREE-RESPONSE QUESTIONS (Form B)

2. A simple random sample of adults living in a suburb of a large city was selected. The age and annual income of each adult in the sample were recorded. The resulting data are summarized in the table below.

	Annual Income			
Age Category	\$25,000-\$35,000	\$35,001-\$50,000	Over \$50,000	Total
21-30	8	15	27	50
31-45	22	32	35	89
46-60	12	14	27	53
Over 60	5	3	7	15
Total	47	64	96	207

(a) What is the probability that a person chosen at random from those in this sample will be in the 31-45 age category?

(b) What is the probability that a person chosen at random from those in this sample whose incomes are over \$50,000 will be in the 31-45 age category? Show your work.

(c) Based on your answers to parts (a) and (b), is annual income independent of age category for those in this sample? Explain.

A) $P(31-45) = \frac{89}{207} \approx 0.42995$

B) $P(31-45 \mid \text{over } 50{,}000) = \frac{35}{96} \approx 0.36458$

C) IF ANNUAL INCOME AND AGE WERE INDEPENDENT, THE ANSWERS IN (A) AND (B) WOULD BE EQUAL. SINCE THEY ARE NOT EQUAL, ANNUAL INCOME + AGE ARE NOT INDEPENDENT FOR MEMBERS OF THIS SAMPLE.

(2003 #3)

3. Men's shirt sizes are determined by their neck sizes. Suppose that men's neck sizes are approximately normally distributed with mean 15.7 inches and standard deviation 0.7 inch. A retailer sells men's shirts in sizes S, M, L, XL, where the shirt sizes are defined in the table below.

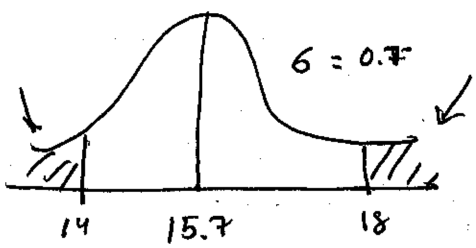
Shirt size	Neck size
S	14 ≤ neck size < 15
M	15 ≤ neck size < 16
L	16 ≤ neck size < 17
XL	17 ≤ neck size < 18

(a) Because the retailer only stocks the sizes listed above, what proportion of customers will find that the retailer does not carry any shirts in their sizes? Show your work.

(b) Using a sketch of a normal curve, illustrate the proportion of men whose shirt size is M. Calculate this proportion.

(c) Of 12 randomly selected customers, what is the probability that exactly 4 will request size M ? Show your work.

A $P(\text{size} \le 14 \text{ or size} \ge 18) = P(\text{size} \le 14) + P(\text{size} \ge 18)$

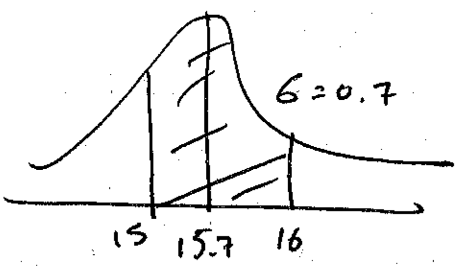


$= 0.007579 + 0.00051$

$= \boxed{0.00809}$

0.007579 0.00051

B



$\boxed{0.50723}$

C Bins? ✓

$P(X=4) = \binom{12}{4}(0.50723)^4(0.4928)^8 = \boxed{0.1139}$

(2001 # 3)

3. Every Monday a local radio station gives coupons away to 50 people who correctly answer a question about a news fact from the previous day's newspaper. The coupons given away are numbered from 1 to 50, with the first person receiving coupon 1, the second person receiving coupon 2, and so on, until all 50 coupons are given away. On the following Saturday, the radio station randomly draws numbers from 1 to 50 and awards cash prizes to the holders of the coupons with these numbers. Numbers continue to be drawn without replacement until the total amount awarded first equals or exceeds $300. If selected, coupons 1 through 5 each have a cash value of $200, coupons 6 through 20 each have a cash value of $100, and coupons 21 through 50 each have a cash value of $50.

(a) Explain how you would conduct a simulation using the random number table provided below to estimate the distribution of the number of prize winners each week.

(b) Perform your simulation 3 times. (That is, run 3 trials of your simulation.) Start at the leftmost digit in the first row of the table and move across. Make your procedure clear so that someone can follow what you did. You must do this by marking directly on or above the table. Report the number of winners in each of your 3 trials.

72749	13347	65030	26128	49067	02904	49953	74674	94617	13317
81638	36566	42709	33717	59943	12027	46547	61303	46699	76423
38449	46438	91579	01907	72146	05764	22400	94490	49833	09258

(A) OBTAIN A 2-DIGIT # FROM RANDOM # TABLE. IF BETWEEN 01-50, USE IT TO REPRESENT THE SELECTED TIX. IGNORE 00, 51-99. DET'M THE AMT OF PRIZE AND STOP WHEN $300 HAS BEEN AWARDED. COUNT # OF WINNERS. IGNORE ANY TIX THAT HAVE BEEN PREVIOUSLY CHOSEN. 19

(B)

~~72~~
~~74~~
~~91~~
33 50 50
47 50 100
~~65~~
03 200 300

3 WINNERS

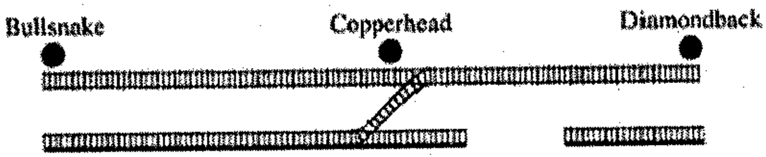
02 200 200
~~61~~ 200
28 50 250
48 50 300

3 WINNERS

06 100 100
~~70~~
~~29~~ 50 150
04 200 350

3 WINNERS.

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.

X = TIME BS → CH
$N(170, 20)$

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.

Y = TIME DB → CH
$N(200, 10)$

These two travel times are independent.

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

(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?

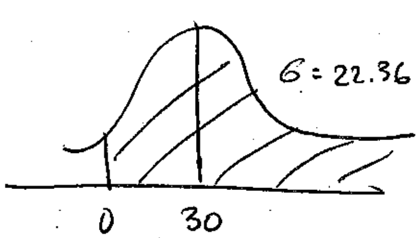
(c) How long should the Snake Gulch Railroad delay the departure of the train from Bullsnake so that the probability that it has to wait is only 0.01 ?

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

$\sigma_{Y-X} = \sqrt{\sigma_x^2 + \sigma_y^2} = \sqrt{10^2 + 20^2} = \sqrt{500} \approx 22.36$

THE DIST. OF Y-X IS NORMAL WITH $\mu_{Y-X} = 30$,
$\sigma_{Y-X} = 22.36$

B WAIT ⇒ Y-X IS POSITIVE



$P(Y-X) \geq 0 = \cancel{0.9082}$
0.9082

∴ THE PROPORTION OF DAYS THE TRAIN WILL HAVE TO WAIT IS 0.91

C LET D = DELAY ⇒ X+D IS TIME WITH DELAY.
X+D IS N(170+D, 20) →

1) LET D BE DELAY NEEDED

X + D IS NORMAL WITH $\mu_{X+D} = 170 + D$, $\sigma_{X+D} = \sigma_X = 20$

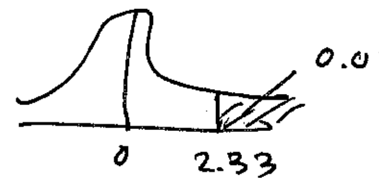
Y IS NORMAL $\mu_Y = 200$ $\sigma_Y = 10$

THE DIFFERENCE $Y - (X + D)$ IS NORMAL

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

$$\sigma_{Y-(X+D)} = \sigma_{Y-X} = 22.36$$

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



TABLE

$$z_{.99} = \frac{0 - \mu_{Y-(X+D)}}{\sigma_{Y-(X+D)}} = \frac{0 - (30 - D)}{22.36} = 2.33$$ TBL

$$(2.33)(22.36) = -30 + D$$

$$52.0988 = -30 + D$$

$$D = \boxed{82.0988 \text{ MIN}}$$

NICE!

(CALC: 2.326479