Statistics – January 2002 Regional Competition

Solutions:

1. $\mu = \frac{\Sigma(scores)}{n} = \frac{1530 + x}{20} = 80 \Rightarrow x = 70.$ **Choice B.**

2. Arrange scores in ascending order: 55,64,67, 68,72,72,78,78,83,83,84,85,85,85,90,92,93,93, 95,98. Median is average of 10^{th} and 11^{th} scores = ½ (83+84)=83.5. **Choice C.**

3. Most repeated score is 85. **Choice D.**

4. The mean of the set is 1620/20=81.
$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{2550}{20}} = \frac{\sqrt{510}}{2}.$$
$A + B = 512 = 2^9 \Rightarrow C = 9.$ **Choice D.**

5. Highest frequency is 6. **Choice C.**

Class	Frequency
50 up to 60	1
60 up to 70	3
70 up to 80	4
80 up to 90	6
90 up to 100	6

6. P(x)=(0.6)(0.03)(0.3) = .54%. Round to 1. **Choice A.**

7. E(x)=(0.65)(0.3)+ (0.35)(-0.05)= 17.75%. **Choice A.**

8. The largest multiple of 7 less than 365 is 52. Since 52x7=364, there will be 52 of every day and 53 Tuesdays (since the year began on a Tuesday). Probability is 52/365. **Choice C.**

9. Separate tables does not matter: 6!=720. **Choice D.**

10. $CV = \frac{\sigma}{\mu} \Rightarrow .7176 = \frac{1.83}{\mu} \Rightarrow \mu = 2.55.$ **Choice C.**

11. Stratified Sampling. **Choice C.**

12. Without the raw data, it is impossible to determine how one additional point will change the equation. **Choice D.**

13. A normal distribution is both bell-shaped and asymptotic, so I and III are true. **Choice E.**

14. Add P(0)+P(1)+P(3) in a binomial distribution.
$$P(0) = \frac{10!}{0!10!}(.3)^0(.7)^{10} = 0.0282475.$$
$$P(1) = \frac{10!}{1!9!}(.3)^1(.7)^9 = 0.1210608.$$
$$P(2) = \frac{10!}{2!8!}(.3)^2(.7)^8 = 0.2334744.$$
Sum=0.383. **Choice C.**

15. $z = \frac{x-\mu}{\sigma} = \frac{28 - 32.68}{2.40} = -1.95.$ Using the area table, P(-1.95<x<0)=0.4744 so P(x<-1.95)=0.5000-.0.4744=.0256. **Choice A.**

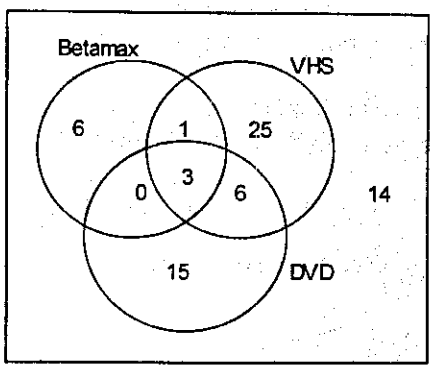
16. P(roll>4)=2/6. P(heads)=1/2.
$$P(roll > 4 \bigcup heads) = \frac{2}{6} + \frac{1}{2} - \left(\frac{2}{6}\right)\left(\frac{1}{2}\right) = \frac{2}{3}.$$
Choice C.

17. Only I is true. **Choice A.**

18. Bayesian Probability Formula.
P(FF)=0.6; P(FF')=0.4;
P(Pizza/FF)=0.9; P(Pizza/FF')=0.25
$$P(FF'/Pizza) = \frac{P(FF')P(Pizza/FF')}{P(FF')P(Pizza/FF') + P(FF)P(Pizza/FF)} =$$
$$\frac{(0.4)(0.25)}{(0.4)(0.25) + (0.6)(0.9)} = \frac{5}{32}.$$ **Choice C.**

Statistics – January 2002 Regional Competition

19. Fourteen have none of the systems. **Choice D.**



20. E(X)=1(3)+0.8(4)+0.5(6)+.25(2)+0.2(13)=12.3 **Choice A.**

21. I and II are unbiased forms of sampling. **Choice A.**

22.

X	Value of the Summation
1	1+i
2	1+i-1=i
3	1+i-1-i=0
4	1+i-1-i+1=1
5	1+i-1-i+1+i=1+i

The table will repeat every fourth value of X, so there is a ¼ chance that the summation is zero. **Choice B.**

23. $u = 3(\frac{x}{2}) + 5(x + \frac{x}{2}) + 9(2x + \frac{x}{2}) + 10(3x + \frac{x}{2}) +$ $7(4x + \frac{x}{2}) + 4(5x + \frac{x}{2}) = 120x.$ $\frac{120x}{38} = 9.47 \Rightarrow x \approx 3.$ **Choice B.**

24. A is the frequency of the point at the 1st quartile: ¼ (38+1)=9.75, implying some frequency between 2x and 3x (or 10 and 15). B is the frequency of the point at the 3rd quartile: ¾ (38+1)=29.25, implying some frequency between 4x and 5x (or 20 and 25). This means the sum A+B is between 30 and 40. **Choice C.**

25. The standard deviation cannot be determined given this information. **Choice E.**

26. The strength of the correlation is dependent upon the magnitude of r. **Choice C.**

27. $39 = (.26)(n+1) \Rightarrow n = 149.$ You must beat all but 15 of the scores: 149-15=134. **Choice A.**

28. Y-int $= \frac{\Sigma Y}{n} - (slope)\frac{\Sigma X}{n} = \frac{33}{7} - \frac{5}{7}\left(\frac{28}{7}\right) = \frac{13}{7}.$ **Choice B.**

29. Confidence Interval $= u \pm z\frac{\sigma}{\sqrt{n}}$. The mean must then be the average of the interval endpoints, 3,300. Therefore, $z\frac{\sigma}{\sqrt{n}} = 35.$ Substituting, $z\left(\frac{240}{\sqrt{144}}\right) = 35 \Rightarrow z = 1.75.$ Using the area table, P(0<x<1.75)=0.4599, so double for confidence interval: 0.9188 or 91.88%. **Choice C.**

30. In Curve M, each individual has an equivalent portion of the total income, so the poorest 1% will own 1% of the income, the poorest 2% will own 2% of the income, and so on. This yields a line whose equation is Y=X. Therefore, A is true, and C and D are true because (0, 0) and (50, 50) are points on the line. Both curves include the point (0, 0) and (100, 100) – because 0% of the population owns no income and 100% of the population owns 100% of the income – so Curves L and M will intersect at least twice. **Choice B.**