PathMD™: Board Review Letter

Author: Philip Ferguson, M.D. **Statistics - Part 2** Volume 1, Number 41

Case #1 The image shown for this question is a Gaussian distribution. The variable X and the variable Y correspond to what percentage of the data represented under the curve?

B. 68% and 95.5%

Answer: B. The area under the curve for a Gaussian distribution between the -1 SD and +1 SD represents 68% of the data, and the area of the curve between the -2 SD and +2 SD represents 95.5% of the data.

Case #2 The image shown for this question is a skewed distribution. The point along the horizontal axis represented by the variable X is best descriptive of which of the following statistical terms?

C. Mode

Answer: C. The mode is defined as the most frequent values occurring in a data distribution. The mean is the average, and the median is the middle value of a numerical series. In a Gaussian distribution, the mean median and mode are the same. When data is skewed, the mean, median, and mode are not equal and therefore the standard deviation is not as meaningful with respect to the distribution of data. In these situations percentile ranks are used to describe the data.

Questions

1. When evaluating a relative operating characteristic (ROC) curve to compare different tests, which of the following will have the greatest efficiency (correctly classify the largest percentage of patients)? Note: sensitivity is plotted from 0 to 100 on the y-axis, and specificity is plotted from 100 to 0 on the x-axis.

   C. The test with the greatest area under the curve

Answer: C. When evaluating an ROC curve, the test with the greatest area under the curve will correctly classify the largest percentage of patients.

2. A test has 95% sensitivity and specificity for a disease with a prevalence of 1 in 500. What is the positive predictive value (PPV) of this test?

   C. 4%

Answer: C. Below is an example of how to set up a 2 x 2 contingency table to easily conceptualize how to calculate the PPV. PPV = TP / TP + FP = 9.5 / 259.5 = 0.04 = 4%

	Test (+)	Test (=)			Test (+)	Test (=)	
Pt. with disease	True (+) TP	False (=) FN	TP+FN	Pt. with disease	9.5	0.5	10
Without disease	False (+) FP	True (=) TN	FP+TN	Without disease	250	4750	5000
	TP + FP	FN + TN			259.5	4750.5	5010

3. A company is introducing a new test for detecting a "test choking" gene. 92 people known to be "test chokers" are tested, and 87 are positive. 610 people who are not "test chokers" are tested, and 576 are negative. Based on this information, calculate the sensitivity and specificity of the new test to detect "test chokers."

    C. Sensitivity – 95%; Specificity – 94%

Answer: C. Below is a 2 x 2 contingency table to help illustrate the calculation of sensitivity and specificity.

	Test (+)	Test (=)			Test (+)	Test (=)	
Pt. with disease	True (+) TP	False (=) FN	TP+FN	Pt. with disease	87	5	92
Without disease	False (+) FP	True (=) TN	FP+TN	Without disease	34	576	610
	TP + FP	FN + TN			121	581	702

Sensitivity = TP / TP + FN = 87 / 87 + 5 = 87 / 92 = 0.95 = 95%

Specificity = TN / TN + FP = 576 / 576 +34 = 576 / 610 = 0.94 = 94%

4. If the prevalence of the "test chokers" is 1 in 750, what is the positive predictive value of the test described in the previous question?

    E. 2%

Answer: E. See explanation below.

	Test (+)	Test (=)			Test (+)	Test (=)	
Pt. with disease	True (+) TP	False (=) FN	TP+FN	Pt. with disease	9.5	0.5	10
Without disease	False (+) FP	True (=) TN	FP+TN	Without disease	450	7050	7500
	TP + FP	FN + TN			459.5	7050.5	7510

PPV = TP / TP + FP = 9.5 / 459.5 = 0.02 = 2%

5. Levy-Jennings plots are useful for evaluating which of the following types of analytical changes within a testing system in sequential plotting of quality control test results?
    A. Trends
    B. Random error
    C. Shifts
    D. A and C are correct
    E. All of the above are correct

Answer: E. Levy-Jennings plots sequentially plot values of quality control specimens, which will detect random error, trends, and shifts.

6. When evaluating a Levy-Jennings plot for random error, shifts, and trends, if present, which would indicate an unstable system?
   A. Trends

Answer: A. Trends occur when a system is gradually changing (often the deterioration of some component of the system – light source, reagents, control material, etc.). Shifts usually occur around a change in reagent lots or maintenance. The machine may need to be recalibrated, or a new mean calculated, but the system is not unstable. Random error as long as it is less than 5% does not represent an unstable system, and is expected. *Test taking strategy.* If you were confused and thought maybe an increase of random error may represent an unstable system, there is no single answer that includes both A and B as correct answers. Therefore, A (trends) is the best answer because it always represents an unstable system in some manner.

7. Positive and negative predictive values are dependent in part of all of the following EXCEPT:

   D. Number of people tested

Answer: D. The number of people tested is not a variable for PPV of NPV.

8. If 10 tests were performed on a "healthy" individual, what is the likelihood all of the results will be within the reference range?

   B. 60%

Answer: B. This question points to the fact that reference ranges by convention exclude 5% of the results (95% confidence interval), which is also approximately 2 standard deviations (95.5%). Considering this, the likelihood of n tests falling within the reference range is = 0.95^n. *Test taking strategy.* This is an easy question if you know the equation and have a scientific calculator. If there is not a scientific available, one can just multiple 0.95 X 0.95 X …. n.

Notes for question set:[1]

[1] PathMD strives for the highest quality and accuracy. However, the *PathMD: Board Review Letter* is for review purposes and not meant for clinical decision making. It should not be used in place of review of primary reference texts and the current medical literature. If inaccuracies are identified, please notify us so that a correction may be published. (info@PathMD.com)

