

Validity Statistics

<u>INSTRUMENT</u>	<u>CRITERION</u>	<u>STATISTIC(S)</u>
numerical	numerical - same scale	<ul style="list-style-type: none"> • calculate discrepancies (Altman & Bland) • describe distribution of discrepancies
numerical	numerical - different scale	<ul style="list-style-type: none"> • correlation [Pearson or Spearman]
numerical	ordered categories (e.g. known groups)	<ul style="list-style-type: none"> • correlation [ranks] • ANOVA [groups]
ordered categories	ordered categories	<ul style="list-style-type: none"> • Kendall's "tau"
numerical or ordered categories	binary	<ul style="list-style-type: none"> • difference in means [parametric/ nonparametric test] • ROC curve
binary	binary	<ul style="list-style-type: none"> • sensitivity (Probability + on instr. if Criterion +) & specificity (Probability – on instr. if Criterion –) • predictive values • (phi) statistic (see Streiner & Norman)

Receiver Operating Characteristic Curve

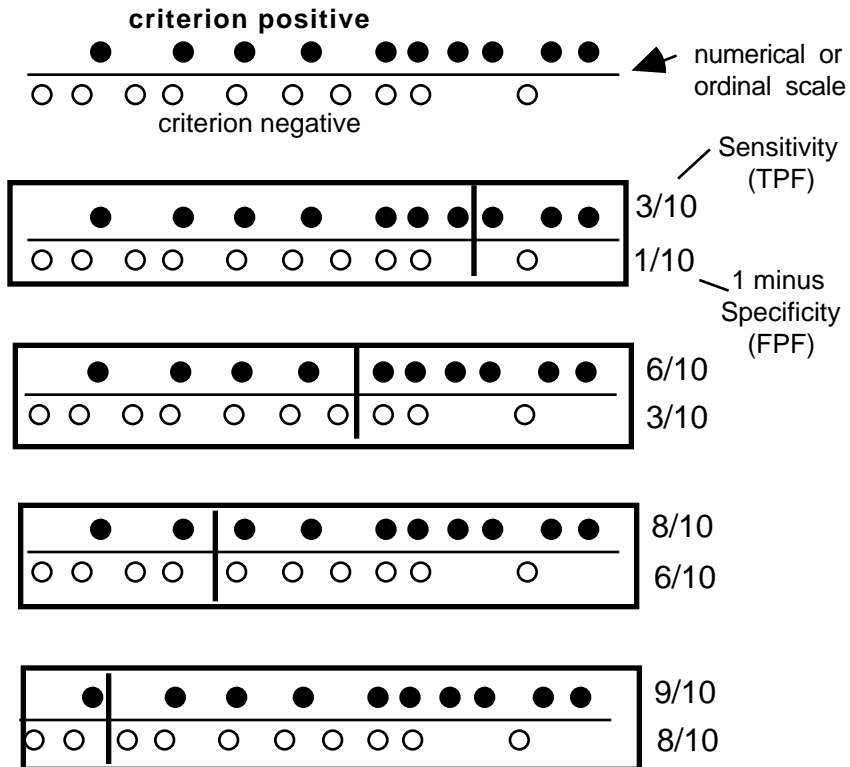
instrument

criterion

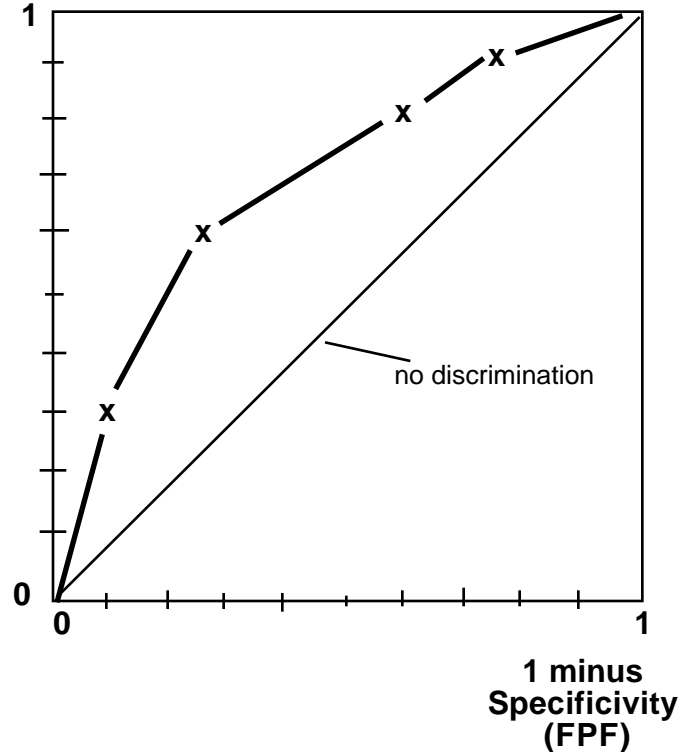
numerical or ordered scale

binary (● or ○)

- SERIES of {sensitivity, specificity} statistics, each based on a different cut-off
- usually plotted on a graph, showing tradeoff between sensitivity and specificity
- Summary statistics (performance)
 - sensitivity at a given (specified) specificity
 - area under the ROC curve



Sensitivity (TPF)



TPF: True Positive Fraction
FPF: False Positive Fraction

Reference: section 5 chapter 13 in 2nd edition of Basic & Clinical Biostatistics by Beth Dawson-Saunders and Robert Trapp, Appleton & Lange, Norwalk (CT)