Types of errors

Jeffrey Leek
Johns Hopkins Bloomberg School of Public Health

Basic terms

In general, **Positive** = identified and **negative** = rejected. Therefore:

True positive = correctly identified

False positive = incorrectly identified

True negative = correctly rejected

False negative = incorrectly rejected

Medical testing example:

True positive = Sick people correctly diagnosed as sick

False positive= Healthy people incorrectly identified as sick

True negative = Healthy people correctly identified as healthy

False negative = Sick people incorrectly identified as healthy.

http://en.wikipedia.org/wiki/Sensitivity_and_specificity

Key quantities

DISEASE

TEST	+	−
+	TP	FP
−	FN	TN

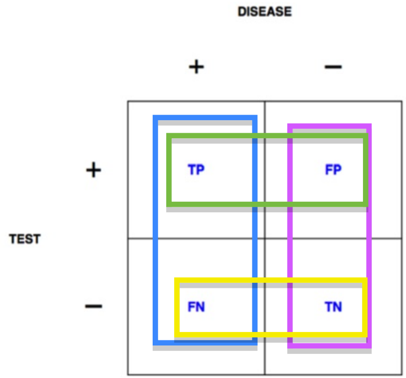
TP + TN = Accuracy

- Sensitivity → Pr (positive test | disease)
- Specificity → Pr (negative test | no disease)
- Positive Predictive Value → Pr (disease | positive test)
- Negative Predictive Value → Pr (no disease | negative test)
- Accuracy → Pr (correct outcome)

http://en.wikipedia.org/wiki/Sensitivity_and_specificity

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

Key quantities as fractions



Sensitivity → TP / (TP+FN)
Specificity → TN / (FP+TN)
Positive Predictive Value → TP / (TP+FP)
Negative Predictive Value → TN / (FN+TN)
Accuracy → (TP+TN) / (TP+FP+FN+TN)

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

Screening tests

Assume that some disease has a 0.1% prevalence in the population. Assume we have a test kit for that disease that works with 99% sensitivity and 99% specificity. What is the probability of a person having the disease given the test result is positive, if we randomly select a subject from

- the general population? also Prevalence = 0.1%
- a high risk sub-population with 10% disease prevalence?

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

General population

		DISEASE	
		+	−
TEST	+	99	999
	−	1	98901

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

General population as fractions

		DISEASE	
		+	−
TEST	+	99	999
	−	1	98901

Sensitivity $\rightarrow$ 99 / (99+1) = 99%

Specificity $\rightarrow$ 98901 / (999+98901) = 99%

Positive Predictive Value $\rightarrow$ 99 / (99+999) $\approx$ 9%

Negative Predictive Value $\rightarrow$ 98901 / (1+98901) $>$ 99.9%

Accuracy $\rightarrow$ (99+98901) / 100000 = 99%

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

At risk subpopulation

		DISEASE +	DISEASE −
TEST	+	9900	900
TEST	−	100	89100

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

At risk subpopulation as fraction

		DISEASE +	DISEASE −
TEST	+	9900	900
TEST	−	100	89100

Sensitivity → 9900 / (9900+100) = 99%

Specificity → 89100 / (900+89100) = 99%

Positive Predictive Value → 9900 / (9900+900) $\approx$ 92%

Negative Predictive Value → 89100 / (100+89100) $\approx$ 99.9%

Accuracy → (9900+89100) / 100000 = 99%

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

Key public health issue

-> false positives

Vast Study Casts Doubts on Value of Mammograms

By GINA KOLATA FEB. 11, 2014

EMAIL
FACEBOOK
TWITTER
SAVE
MORE

One of the largest and most meticulous studies of mammography ever done, involving 90,000 women and lasting a quarter-century, has added powerful new doubts about the value of the screening test for women of any age.

It found that the death rates from breast cancer and from all causes were the same in women who got mammograms and those who did not. And the screening had harms: One in five cancers found with mammography and treated was not a threat to the woman's health and did not need treatment such as chemotherapy, surgery or radiation.

The study, published Tuesday in The British Medical Journal, is one of the few rigorous

Nearly 75 percent of American women 40 and over say they had a mammogram in the past year. Damian Dovarganes/Associated Press

http://www.biostat.jhsph.edu/~iruczins/teaching/140.615/

Key public health issue

The New York Times

Business Day

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | AR

Search | Global | DealBook | Markets | Economy | Energy | Media | Technology

Looser Guidelines Issued on Prostate Screening

By ANDREW POLLACK
Published: May 3, 2013

In a major shift, the American Urological Association has pulled back its strong support of prostate cancer screening, saying that the testing should be considered primarily by men aged 55 to 69.

The association had staunchly defended the benefits of screening men with the prostate test, even after a government advisory committee, the United States Preventive Services Task Force, said in 2011 that healthy men should not be screened because far more men would be harmed by unnecessary prostate cancer treatments than would be saved from death.

FACEBOOK
TWITTER
GOOGLE+
SAVE
EMAIL
SHARE
PRINT
REPRINTS

For continuous data

Mean squared error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(\text{Prediction}_i - \text{Truth}_i)^2$$

Root mean squared error (RMSE):

`to get the same unit:`

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{Prediction}_i - \text{Truth}_i)^2}$$

Common error measures

1. Mean squared error (or root mean squared error)
    - Continuous data, sensitive to outliers
2. Median absolute deviation
    - Continuous data, often more robust
3. Sensitivity (recall)
    - If you want few missed positives
4. Specificity
    - If you want few negatives called positives
5. Accuracy
    - Weights false positives/negatives equally
6. Concordance
    - One example is kappa