ECE531 Screencast 8.7: Neyman-Pearson Hypothesis Testing Example

D. Richard Brown III

Worcester Polytechnic Institute

Example: 10 Coin Flips

Coin flipping problem with a probability of heads of either $q_0 = 0.5$ or $q_1 = 0.8$. We observe ten flips of the coin and count the number of heads.

$$P = \begin{bmatrix} 0.0010 & 0.0000 \\ 0.0098 & 0.0000 \\ 0.0439 & 0.0001 \\ 0.1172 & 0.0008 \\ 0.2051 & 0.0055 \\ 0.2461 & 0.0264 \\ 0.2051 & 0.0881 \\ 0.1172 & 0.2013 \\ 0.0439 & 0.3020 \\ 0.0098 & 0.2684 \\ 0.0010 & 0.1074 \end{bmatrix} \text{ and } L = \begin{bmatrix} 0.0001 \\ 0.0004 \\ 0.0017 \\ 0.0067 \\ 0.0268 \\ 0.1074 \\ 0.4295 \\ 1.7180 \\ 6.8719 \\ 27.4878 \\ 109.9512 \end{bmatrix}$$

What is ρ^{NP} and $\beta = P_D(\rho^{\text{NP}})$ when $\alpha = 0.001$, $\alpha = 0.01$, $\alpha = 0.1$?

Detailed Calculations for $\alpha = 0.1$ Case (part 1 of 2)

$L_{\text{sorted}} = [109.9512, 27.4878, 6.8719, 1.7180, 0.4295 \ldots]$.

likelihood threshold v	δ^v	$P_{\text{fp}}(\delta^v) = \sum_{\ell:L_\ell > v} P_{\ell,0}$
109.9512	decide $\mathcal{H}_1$ if $y > 10$	0.0000
27.4878	decide $\mathcal{H}_1$ if $y > 9$	0.0010
6.8719	decide $\mathcal{H}_1$ if $y > 8$	0.0108
1.7180	decide $\mathcal{H}_1$ if $y > 7$	0.0547
0.4295	decide $\mathcal{H}_1$ if $y > 6$	0.1719

Hence, $v = 1.17180$ is the minimum value such that $\sum_{\ell:L_\ell > v} P_{\ell,0} \le \alpha$.

Now solve for the randomization γ:

$$\gamma = \frac{\alpha - \sum_{\ell:L_\ell > v} P_{\ell,0}}{\sum_{\ell:L_\ell = v} P_{\ell,0}} = \frac{0.1 - 0.0547}{0.1172} = 0.3865$$

Detailed Calculations for $\alpha = 0.1$ Case (part 2 of 2)

Hence, the Neyman-Pearson criterion decision rule for the $\alpha = 0.1$ case is given as

$$\rho^{\text{NP}}(y) = \begin{cases} 1 & \text{if } L_\ell > v \\ \gamma & \text{if } L_\ell = v \\ 0 & \text{if } L_\ell < v \end{cases}$$

with $v = 1.7180$ and $\gamma = 0.3865$. This is more directly written as

$$\rho^{\text{NP}}(y) = \begin{cases} 1 & \text{if } y > 7 \\ 0.3865 & \text{if } y = 7 \\ 0 & \text{if } y < 7 \end{cases}$$

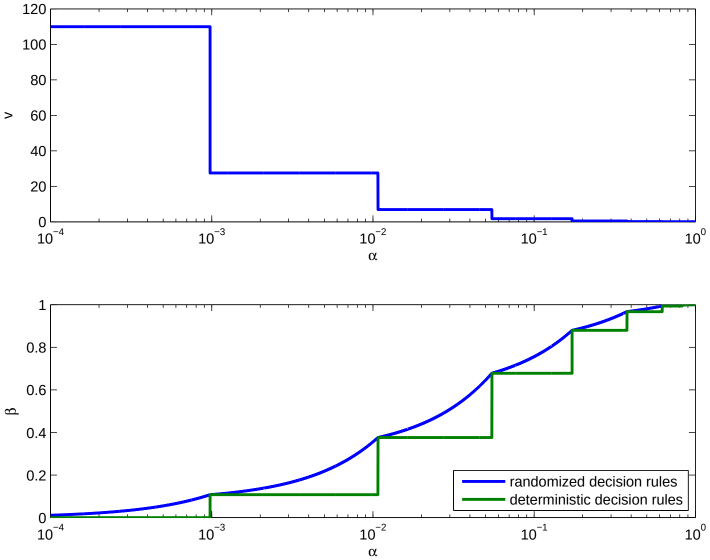
Check the false positive probability:

$$P_{\text{fp}}(\rho^{\text{NP}}) = 0.0010 + 0.0098 + 0.0439 + \gamma \cdot 0.1172 = 0.1000 = \alpha$$

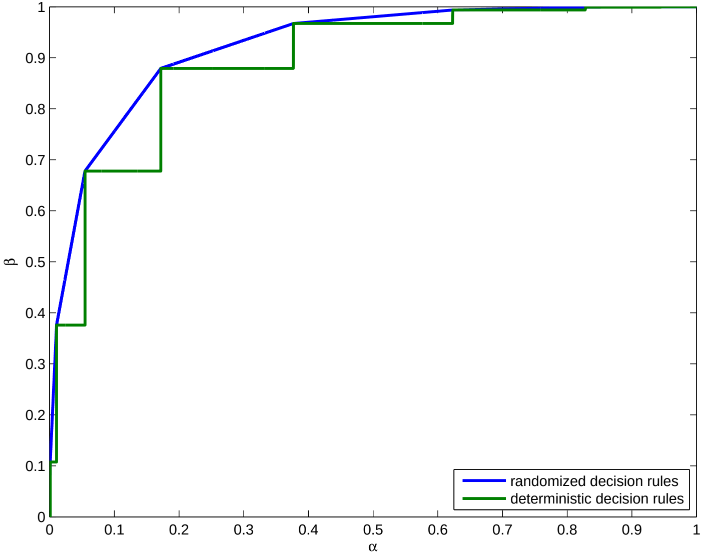
Compute the probability of detection:

$$P_{\text{D}}(\rho^{\text{NP}}) = 0.1074 + 0.2684 + 0.3020 + \gamma \cdot 0.2013 = 0.7556 = \beta$$

Example: Randomized vs. Deterministic Decision Rules

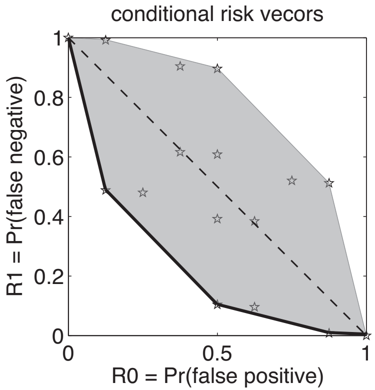


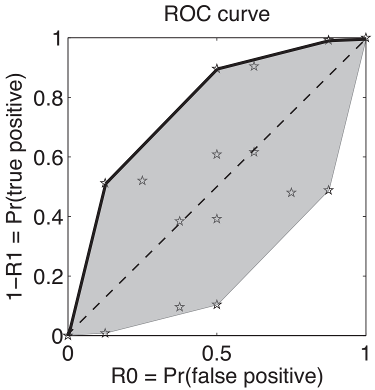
Example: Same Results Except Linear Scale



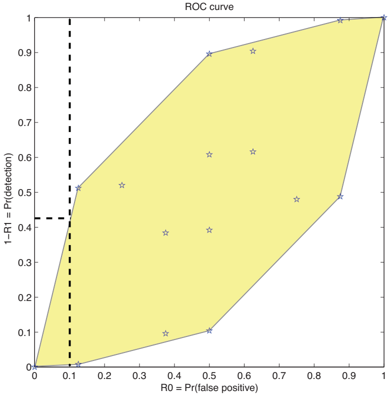
Remarks: 1 of 2

The blue line on the previous slide is called the **Receiver Operating Characteristic** (ROC). An ROC plot shows the probability of detection $\beta = P_D = 1 - R_1$ as a function of $\alpha = R_0$. The ROC plot is directly related to our conditional risk vector plot.





Remarks: 2 of 2



The N-P criterion seeks a decision rule that **maximizes the probability of detection** subject to the constraint that the probability of false alarm must be no greater than α.

$$\rho^{\text{NP}} = \arg\max_{\rho} P_D(\rho)$$
$$\text{s.t. } P_{\text{fp}}(\rho) \leq \alpha$$

- The term **power** is often used instead of "probability of detection". The N-P decision rule is sometimes called the "most powerful test of significance level α".
- Intuitively, the power of a test should increase with the significance level of the test.