# Accuracy with latent categorical variable

Mc

May 27, 2014

## Introduction

In plots we recast the results of simulations in terms of accuracy. We compute the accuracy of each method (continous or categorical), for each level of $\rho$ (see below) by computing the the following quantities:

- false positive (FP) runs with with a significant test under a true null hypothesis
- true positive (TP) runs with a significant test under a false null-hypothesis
- true negative (TN) runs with a nonsignificant result under a true null-hypothesis
- false negative (FN) runs with a nonsignificant result under a false null-hypothesis

Plots to be produced:

- Sensitivity for all 4 of the decision possibilities (continuous ignoring categorical, categorical ignoring continuous, both, either), with X axis being rho (Figure 1)
- PPV for the 4 decision possibilities, with X axis being rho (Figure 2)
- Bar chart with the specificity for the 4 decision possibilities
- Bar chart with the NPV (aggregated over rho) for the 4 decision possibilities

## Setup

### Model

A categorical latent variable ( $\xi$ ) and a continuous one ($\eta$ are created with N cases, sharing a correlation equal to $\rho$. A measure $x$ of $\xi$ is created with reliability $rel$, and then is dichotomized accordingly to $p$ $1-p$ into $c$. The correlations $r_pe = r(\eta, x)$ and $r_pb = r(\eta, c)$ are computed, their p-value and significance (at .05) is recorded.

## Design

$\rho = (0, .1, .2, .3, .4, .5, .6, .7)$ $rel = (0.3, 0.4, 0.5, 0.6, 0.7, 0.80.9)$

## Computation of quantities

- Continuous false positive (FP_C) freq of runs with continuous test p.<.05 and $\rho$=0
- Continuous true positive (TP_C) freq of runs with continuous test p.<.05 and $\rho$ >0
- Continuous true negative (TN_C) freq of runs with continuous test p.>=.05 and $\rho$=0
- false negative (FN_C) freq of runs with continuous test p.>=.05 and $\rho$>0
- PPV is defined as TP/(TP+FP)
- NPV is defined as TN/(TN+FN)

The same quantities are computed for the categorical indicator (*_S).

```
Accuracy for continuous indicator

  rho     SENS_C    SPEC_C       PPV       NPV
1 0.1 0.09957143 0.9497143 0.6644423 0.5133194
2 0.2 0.25342857 0.9497143 0.8344309 0.5598787
3 0.3 0.44314286 0.9497143 0.8980892 0.6303812
4 0.4 0.60228571 0.9497143 0.9229422 0.7048346
5 0.5 0.72042857 0.9497143 0.9347544 0.7725741
6 0.6 0.80942857 0.9497143 0.9415088 0.8328740
7 0.7 0.86457143 0.9497143 0.9450344 0.8751975

Accuracy for categorical indicator

  rho     SENS_S    SPEC_S       PPV       NPV
1 0.1 0.09857143 0.9491429 0.6596558 0.5128918
2 0.2 0.23714286 0.9491429 0.8234127 0.5544059
3 0.3 0.40400000 0.9491429 0.8881910 0.6142751
4 0.4 0.55128571 0.9491429 0.9155397 0.6789985
5 0.5 0.66071429 0.9491429 0.9285284 0.7366670
6 0.6 0.74485714 0.9491429 0.9360862 0.7881376
7 0.7 0.80385714 0.9491429 0.9404981 0.8287389

Accuracy for BOTH indicators significant
```

```
  rho     SENS_B    SPEC_B       PPV       NPV
1 0.1 0.06114286 0.9771429 0.7278912 0.5099911
2 0.2 0.17800000 0.9771429 0.8862020 0.5431158
3 0.3 0.35200000 0.9771429 0.9390244 0.6012658
4 0.4 0.50814286 0.9771429 0.9569545 0.6651755
5 0.5 0.62814286 0.9771429 0.9648892 0.7243461
6 0.6 0.72214286 0.9771429 0.9693193 0.7785999
7 0.7 0.78971429 0.9771429 0.9718706 0.8229066

Accuracy for EITHER indicators significant

  rho    SENS_E    SPEC_E       PPV       NPV
1 0.1 0.1370000 0.9217143 0.6363636 0.5164492
2 0.2 0.3125714 0.9217143 0.7997076 0.5727983
3 0.3 0.4951429 0.9217143 0.8634778 0.6461045
4 0.4 0.6454286 0.9217143 0.8918279 0.7221849
5 0.5 0.7530000 0.9217143 0.9058257 0.7886566
6 0.6 0.8321429 0.9217143 0.9140122 0.8459420
7 0.7 0.8787143 0.9217143 0.9181967 0.8837146
```

**Figure 1: Sensitivity for all 4 of the decision possibilities (continuous ignoring categorical, categorical ignoring continuous, both, either), with X axis being rho**

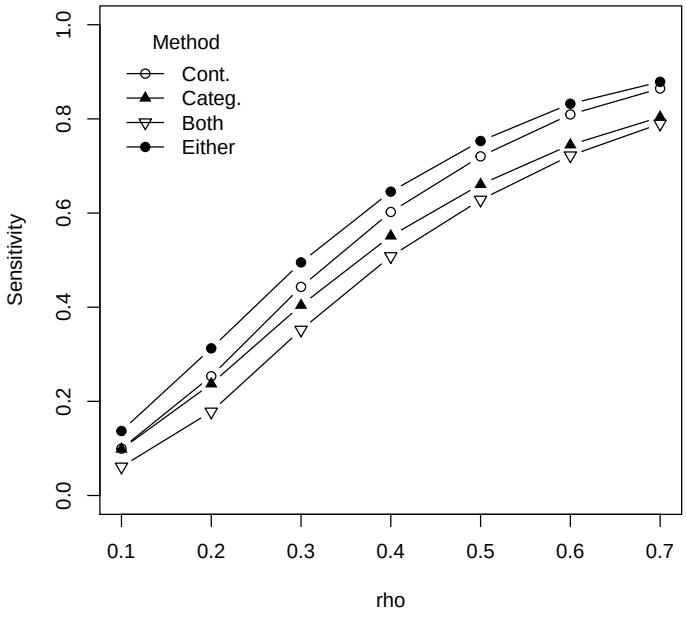

**Figure 2: PPV for the 4 decision possibilities, with X axis being rho**

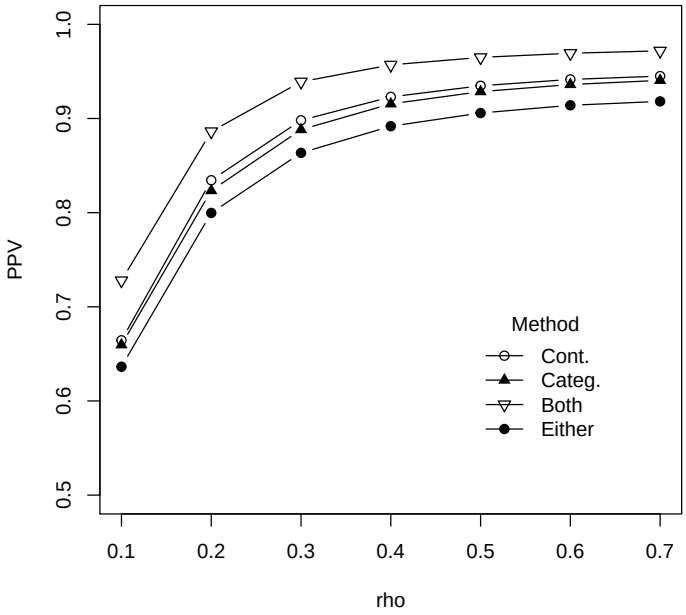

**Figure 3: Bar chart with the specificity for the 4 decision possibilities**

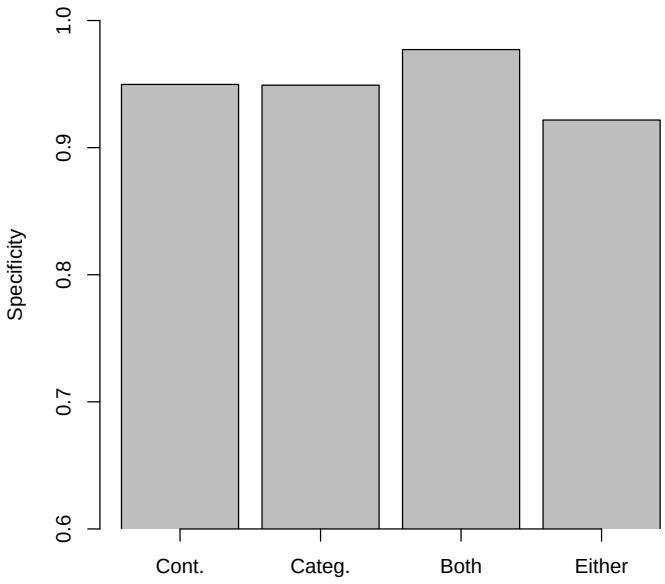

**Figure 4: NPV for the 4 decision possibilities, with X axis being rho**