

Supplementary Material

Classification of Trends in Dose-Response Microarray Experiments Using Information Theory Selection Methods

D. Lin^{*1}, Z. Shkedy¹, T. Burzykowski¹, M. Aerts¹, H. W. H. Gohlmann², A. De Bondt², T. Perera², T. Geerts², I. Van den Wyngaert² and L. Bijnen²

¹Hasselt University, I-Biostat, Agoralaan 1, Building D, 3590, Deipenbeek, Belgium

²Johnson & Johnson PR&D, Turnhoutseweg 30, 2340, Beerse, Belgium

1. EXAMPLE OF OBTAINING TEST STATISTICS AND PERMUTATION P-VALUES FOR ONE PROBE SET

We use one probe set generated from the following codes to illustrate how to obtain the five test statistics and permutation raw p-value for both directions. The analysis is implemented by using the R library IsoGene (available at <http://cran.r-project.org/web/packages/IsoGene/index.html>).

```
> library(IsoGene)
```

```
> load("dataexample.RData")
```

A printout of the first 10 lines is given below.

```
> data [1:10,1:6]
```

	X1	X1.1	X1.2	X2	X2.1	X2.2
g1	6.923109	7.024719	7.170328	7.219297	7.076908	7.404949
g2	5.107275	5.092935	5.255918	5.312913	4.893855	4.596591
g3	5.913526	6.026197	5.141728	5.828770	5.269202	5.461664
g4	4.919469	4.908159	3.500307	4.814068	4.139949	4.278321
g5	6.002091	5.878718	5.777668	6.214799	5.895586	6.163291
g6	7.162715	7.294693	6.903935	7.223069	6.972928	7.412160
g7	4.049696	4.748409	3.845498	4.780287	4.076589	4.300242
g8	3.191931	4.326571	3.771206	3.570291	2.179324	3.988911
g9	6.487708	6.285804	6.229814	6.109103	6.340837	5.931840
g10	6.695870	6.687039	6.652153	6.503670	6.387794	6.698711

```
> x <- c(rep(1, 3), rep(2, 3), rep(3, 3), rep(4, 3))
```

```
> gene1 <- as.matrix(exampleData[2, ])
```

```
> par(mfrow = c(1, 2))
```

```
> IsoPlot(x, y = gene1)
```

```
> IsoPlot(x, y = gene1, type = "ordinal", add.curve = TRUE)
```

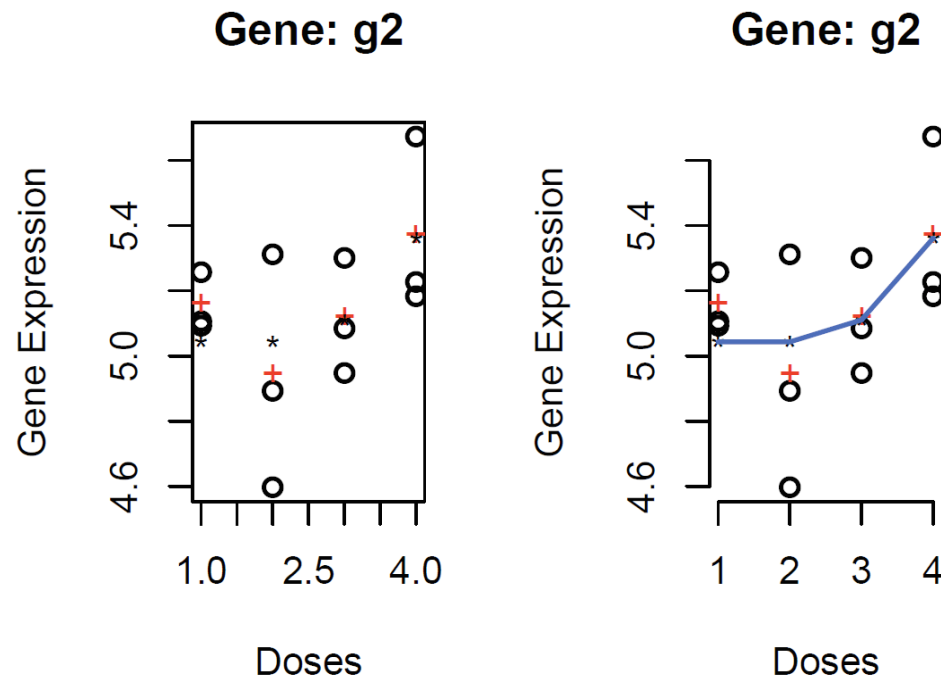


Fig. (S1). The data points are plotted as circles, while sample means as pluses. The right panel additionally plots the fitted increasing isotonic regression model (blue solid line).

The left panel in Fig. (S1) shows the original data points (as circles) and sample means (as pluses) for each dose. The right panel in Fig. (S1) shows the increasing isotonic regression model (blue solid line) fitted on the data. The fitted monotonic line does not indicate the significance of the test, but simply shows a more likely increasing (or decreasing) trend.

```
stat1 <- IsoGene1(x, gene1)
```

The object stat1 contains the information about the five test statistics and the direction for which the likelihood is maximized.

```
> stat1
```

```
$E2.up
```

```
[1] 0.2697894
```

```
$Williams.up
```

```
[1] 1.040134
```

```
$Marcus.up
```

```
[1] 1.581191
```

```
$M.up
```

```
[1] 1.205802
```

```
$ModM.up
```

```
[1] 1.278946
```

```
$E2.dn
```

```
[1] 0.0008106545
```

```
$Williams.dn
```

```
[1] -0.08238646
```

```
$Marcus.dn
```

```
[1] -0.08238646
```

```
$M.dn
```

```
[1] -0.05370908
```

```
$ModM.dn
```

```
[1] -0.06004858
```

```
$direction
```

```
[1] "u"
```

For a single gene, the function `IsopvaluePlot()` can be used to show the p^{Up} and p^{Down} -values for a given test statistic:

```
> IsopvaluePlot(x, y, niter, stat = c("E2", "Williams", "Marcus", "M", "ModifM"))
```

We use one gene as an example to illustrate how p^{Up} and p^{Down} -values (in the upper and lower panels of Fig. S2) are obtained. In Fig. (S2), the observed test statistics are drawn as the dashed line, and the values of the test statistics obtained from permutations are spread over the x-axis. For this gene, the p^{Up} is much smaller as compared to the p^{Down} since $T^{\text{Up}} \gg T^{\text{Down}}$, which implies a possible increasing trend in the data.

```
> IsopvaluePlot(x, gene1, niter = 1000, stat = "E2")
```

For more details on how to use the R package IsoGene, we refer to Pramara *et al.* (2009a), Pramara *et al.* (2009b).

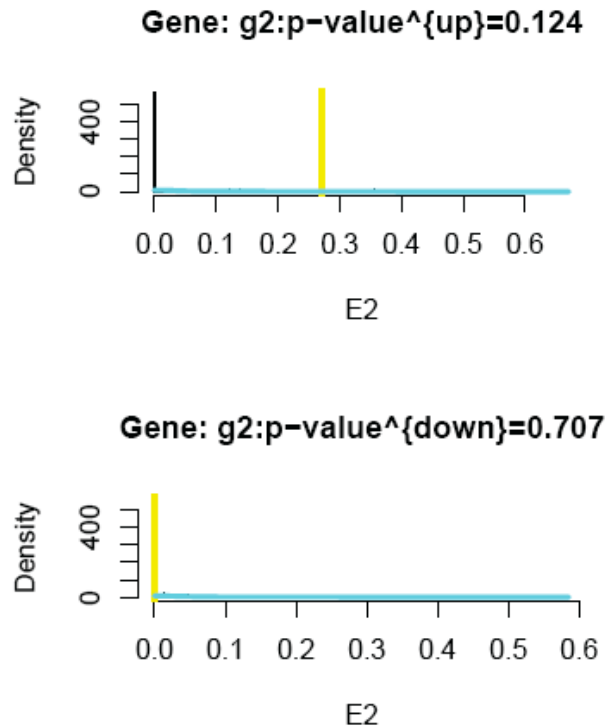


Fig. (S2). The p^{Up} and p^{Down} -values using $\bar{E}_{01}^2$ for an example gene. The dashed line is the observed test statistic value. In the upper panel, the dashed line (at the right) is larger than most of the test statistics from permutations, which results in a small p^{Up} -value. In the lower panel, the dashed line (close to zero) is smaller than most of the test statistics from permutations, which results in a large p^{Down} -value.

REFERENCES

- Pramana S., Lin D., Haldermans P., Shkedy Z., Verbeke T., GÄohlmann H. W.H., De Bondt A., Talloen W., and Bijns L.. (2009) Analysis of Dose-response Studies in Microarray Experiments Using the R IsoGene Package (*submitted to Journal of Statistical Software*).
- Pramana S., Lin D., Haldermans P., Shkedy Z., Verbeke T., GÄohlmann H. W.H., De Bondt A., Talloen W., and Bijns L.. (2009) IsoGene: An R Package for Analyzing Dose-response Studies in Microarray Experiments. R Journal (*under revision*).