

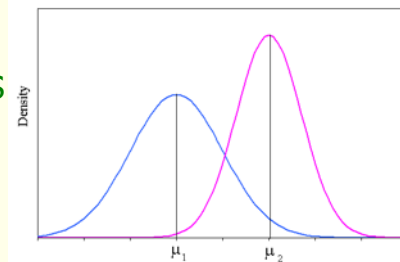
TWO-SAMPLE COMPARISONS

⇒ GENERAL OVERVIEW OF TWO-SAMPLE TESTS

⇒ EXAMPLES OF MICROARRAY EXPERIMENTS

Single-slide methods
Two-slide experiments
Replicate slides

⇒ ALTERNATIVE APPROACHES
AND ADDITIONAL TOPICS



TWO-SAMPLE TESTS *

	Parametric (Normal assumptions)	Nonparametric
Independent samples	t -test	Mann-Whitney (or Wilcoxon-Mann-Whitney)
Related samples	Paired t -test	Wilcoxon

* Here we are interested on comparing the location of two distributions (such as mean, median, ranks, etc.). In other situations we may want to compare two proportions, variances, skewness, etc.

TWO-SAMPLE TESTS

Note:

- ① There are some alternative approaches between Gaussian-based and nonparametric methods, such as the generalized linear models (McCullagh and Nelder, 1989; Tempelman, 1998) or methods employing thick-tailed (Liu, 1996; Stranden and Gianola, 1999; Rosa, 1999) and/or skewed distributions (Fernandez and Steel, 1998; von Rohr and Ina, 2002), among others.
- ② The sampling-based methods (such as the Bootstrap and Permutation test) are useful and flexible alternatives to the nonparametric methods as well (Efron and Tibshirani, 1998). To be discussed later.

THE TWO-SAMPLE *t*-TEST

➔ The case of equal variances

Clone i

Sample 1	Sample 2
y_{11}	y_{21}
y_{12}	y_{22}
...	...
\bar{y}_{1n1}	\bar{y}_{2n2}
$\bar{y}_1 \pm s_1$	$\bar{y}_2 \pm s_2$

Settings: $\left\{ \begin{array}{l} y_{1j} \sim N(\mu_1, \sigma_1^2) \text{ and } y_{2j} \sim N(\mu_2, \sigma_2^2) \\ H_0 : \mu_1 = \mu_2 \text{ vs. } H_1 : \mu_1 \neq \mu_2 \\ \sigma_1^2 = \sigma_2^2 \end{array} \right.$

Test statistic:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t_{(n_1+n_2-2)}$$

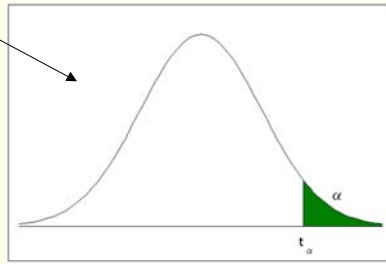
$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

'weighted' average of s_1^2 and s_2^2

CRITICAL REGIONS AND P-VALUES

If $t_i > t_\alpha \Rightarrow$ Gene i differentially expressed

$$t_i = \frac{\bar{y}_{i1} - \bar{y}_{i2}}{se_i}$$



Note: If two-sided test, compare $|t_i|$ with $t_{\alpha/2}$.

THE TWO-SAMPLE t -TEST

COMMENT: This t -test is equivalent to a one-way two-sample ANOVA (F test)

S.V.	D.F.	S.S.	M.S.	F
Samples	1	$\sum_i n_i \bar{y}_i^2 - C$	$\frac{\sum_i n_i \bar{y}_i^2 - C}{1}$	$\frac{\sum_i n_i \bar{y}_i^2 - C}{n_1 + n_2 - 2}$
Residual	$n_1 + n_2 - 2$	$\sum_j \sum_i y_{ij}^2 - \sum_i n_i \bar{y}_i^2$	$\frac{\sum_j \sum_i y_{ij}^2 - \sum_i n_i \bar{y}_i^2}{n_1 + n_2 - 2}$	---
Total	$n_1 + n_2 - 1$	$\sum_j \sum_i y_{ij}^2 - C$	---	---

$$F_{(1,\phi)} = t_\phi^2$$

$$C = \frac{(n_1 \bar{y}_1 + n_2 \bar{y}_2)^2}{n_1 + n_2} \quad (\text{"correction"})$$

THE TWO-SAMPLE t -TEST

COMMENT: If $\sigma_1^2 \neq \sigma_2^2$, the t -test becomes:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{(\varphi)}, \text{ where: } \varphi = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 + 1}}$$

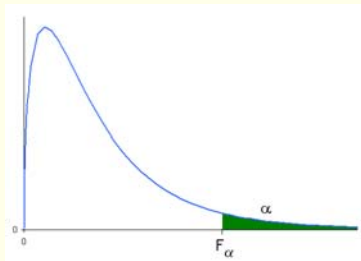
(Satterthwaite, 1946)

THE TWO-SAMPLE t -TEST

COMMENTS: Testing for equal variances.

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ vs. } H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$F = \frac{\max[s_1^2, s_2^2]}{\min[s_1^2, s_2^2]} \sim F_{(\text{d.f. numerator; d.f. denominator})}$$



EXAMPLE and SAS CODE

T1	T2
5.71	5.44
5.57	5.49
5.62	5.48
5.91	5.77
5.94	5.82
5.74	---

```

data t;
input trt $ y;
cards;
T1 5.71
T1 5.57
T1 5.62
T1 5.91
T1 5.94
T1 5.74
T2 5.44
T2 5.49
T2 5.48
T2 5.77
T2 5.82
;
run;

proc print data=t;
run;

proc ttest alpha=.01 data=t;
class trt;
var y;
run;
    
```

Alternative: ANALYST

RESULTS

Obs	trt	y
1	T1	5.71
2	T1	5.57
3	T1	5.62
4	T1	5.91
5	T1	5.94
6	T1	5.74
7	T2	5.44
8	T2	5.49
9	T2	5.48
10	T2	5.77
11	T2	5.82

ANALYST				
Sample Statistics				
Group	N	Mean	Std. Dev.	Std. Error
T1	6	5.748333	0.1501	0.0613
T2	5	5.6	0.1799	0.0804

Hypothesis Test				
Null hypothesis:		Mean 1 - Mean 2 = 0		
Alternative:		Mean 1 - Mean 2 \neq 0		
If Variances Are	t statistic	Df	Pr > t	
Equal	1.494	9	0.1695	
Not Equal	1.467	7.87	0.1812	

PROC TTEST					
T-Tests					
Variable	Method	Variances	DF	t Value	Pr > t
y	Pooled	Equal	9	1.49	0.1695
y	Satterthwaite	Unequal	7.87	1.47	0.1812

Equality of Variances					
Variable	Method	Num DF	Den DF	F Value	Pr > F
y	Folded F	4	5	1.44	0.6907

THE PAIRED *t*-TEST

➔ before/after; left/right; twins, etc.

Subject	T1	T2
1	Y_{11}	Y_{21}
2	Y_{12}	Y_{22}
...
n	Y_{1n}	Y_{2n}



Subject	Diference (T1 - T2)
1	d_1
2	d_2
...	...
n	d_n
mean ± std	$\bar{d} \pm S_d$

$H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$

$H_0 : d = 0$ vs. $H_1 : d \neq 0$

Test statistic:

$$t = \frac{\bar{d}}{\sqrt{\frac{S_d^2}{n}}} \sim t_{(n-1)}$$

EXAMPLE and SAS CODE

Subject	Before	After
1	54	56
2	61	65
3	50	52
4	74	73
5	80	82
6	62	61
7	52	59
8	60	64

```

data paired;
input subj before after;
cards;
1 54 56
2 61 65
3 50 52
4 74 73
5 80 82
6 62 61
7 52 59
8 60 64
;

proc print data=paired noobs;
title 'Two related samples';
run;

proc ttest data=paired alpha=.01;
paired before*after;
run;
    
```

Alternative: ANALYST

RESULTS

subj	before	after
1	54	56
2	61	65
3	50	52
4	74	73
5	80	82
6	62	61
7	52	59
8	60	64

ANALYST

Sample Statistics				
Group	N	Mean	Std. Dev.	Std. Error
before	8	61.625	10.555	3.7318
after	8	64	9.6214	3.4017

Hypothesis Test

Null hypothesis: Mean of (before - after) = 0
 Alternative: Mean of (before - after) \neq 0

t Statistic	Df	Prob > t
-2.517	7	0.0400

PROC TTEST

Statistics										
Difference	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err	Minimum	Maximum
before - after	8	-5.678	-2.375	0.9276	1.5683	2.6693	7.1005	0.9437	-7	1

T-Tests			
Difference	DF	t Value	Pr > t
before - after	7	-2.52	0.0400

THE PAIRED *t*-TEST

COMMENT: This *t*-test is equivalent to a two-sample RBD ANOVA (F test)

S.V.	D.F.	S.S.	M.S.	F
Pairs	n-1	$\sum_j 2\bar{y}_j^2 - C$	$\frac{\sum_j 2\bar{y}_j^2 - C}{n-1}$	---
Samples	1	$\sum_i n_i \bar{y}_i^2 - C$	① $\sum_i n_i \bar{y}_i^2 - C$	①/②
Residual	n-1	$\sum_i \sum_j y_{ij}^2 - \sum_i n_i \bar{y}_i^2 - \sum_j 2\bar{y}_j^2 + C$	② $\frac{\text{SS Residual}}{n-1}$	---
Total	2n-1	$\sum_i \sum_j y_{ij}^2 - C$	---	---

THE PAIRED t -TEST

COMMENT:

If you wish to test: $H_0: d = c$ vs. $H_1: d \neq c$

Use:

$$t = \frac{\bar{d} - c}{\sqrt{\frac{s_d^2}{n}}} \sim t_{(n-1)}$$

NONPARAMETRIC METHODS

Note: The t -tests (paired or unpaired; homogeneous or heterogeneous variances) are based on normality assumptions. If the data are not well approximated by a normal distribution, the t -test may not be appropriate.

THE MANN-WHITNEY TEST

Sample 1	Sample 2
Y_{11}	Y_{21}
Y_{12}	Y_{22}
...	...
Y_{1n_1}	Y_{1n_2}

Sort the $n = n_1 + n_2$
values

Replace y_{ij} by its
overall rank r_{ij}

Sample 1	Sample 2
3 (r_{11})	5 (r_{21})
1 (r_{12})	2 (r_{22})
...	...
6 (r_{1n_1})	9 (r_{2n_2})

$$R_1 = \sum_{j=1}^{n_1} r_{1j} \quad \text{and} \quad R_2 = \sum_{j=1}^{n_2} r_{2j} \quad (\text{sums of ranks})$$

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad \left(\text{or } U' = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 ; U = n_1 n_2 - U' \right)$$

(P-values from appropriate tables)

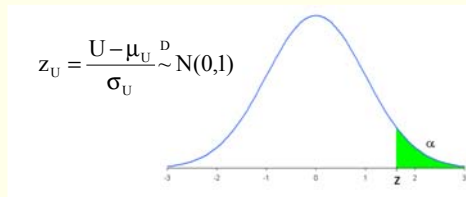
THE MANN-WHITNEY TEST

(Normal approximation for U)

$$U \stackrel{D}{\sim} N(\mu_U, \sigma_U^2)$$

$$\mu_U = \frac{n_1 n_2}{2}$$

$$\sigma_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$



P-value = α (one-sided)

or

P-value = 2α (two-sided)

EXAMPLE and SAS CODE

Male	Female
76	68
73	69
74	64
72	65
67	66
70	---
71	---

Alternative: ANALYST

```
data mn;
input group $ height;
cards;
male 76
male 73
male 74
male 72
male 67
male 70
male 71
female 68
female 69
female 64
female 65
female 66
;
run;

proc print data=mn;
title 'Wilcoxon-Mann-Whitney';
run;

proc npar1way data=mn wilcoxon;
class group;
var height;
run;
```

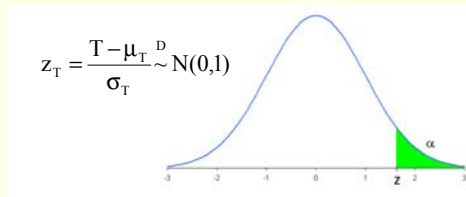

THE WILCOXON TEST

(Normal approximation for T)

$$T^D \sim N(\mu_T, \sigma_T^2)$$

$$\mu_T = \frac{n(n+1)}{4}$$

$$\sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$$



P-value = α (one-sided)

or

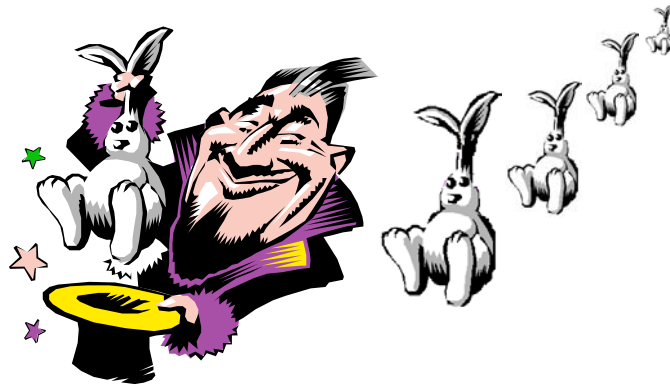
P-value = 2α (two-sided)

EXAMPLE

Deer	Hindleg	Foreleg	Difference	Rank	Signed
1	142	138	4	4.5	4.5
2	140	136	4	4.5	4.5
3	144	147	-3	3	-3
4	144	139	5	7	7
5	142	143	-1	1	-1
6	146	141	5	7	7
7	149	143	6	9.5	9.5
8	150	145	5	7	7
9	142	136	6	9.5	9.5
10	148	146	2	2	2

$$T = 4.0 \text{ (} P < 0.01 \text{)}$$

SAMPLING-BASED TECHNIQUES



THE PERMUTATION TEST

- ➔ The basic idea is attractively simple and free of mathematical assumptions
- ➔ There is a close connection with Bootstrap (to be discussed later) case of equal variances

Suppose:

Experiment	
Trt 1	Trt 2
Y_{11}	Y_{21}
Y_{12}	Y_{22}
...	...
Y_{1n1}	Y_{2n2}
$\bar{Y}_1 \pm s_1$	$\bar{Y}_2 \pm s_2$

From distribution F

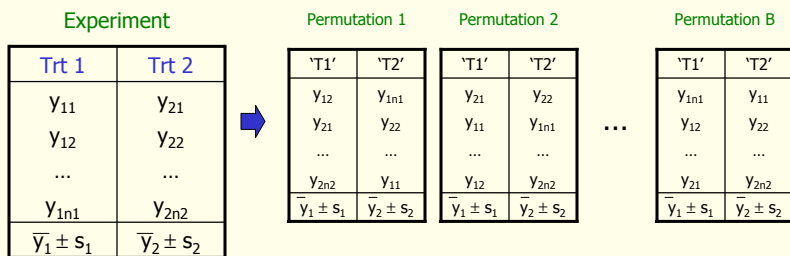
From distribution G

$H_0 : F = G$ vs. $H_1 : F \neq G$

THE PERMUTATION TEST

- ① Define the statistics (e.g. $t = \frac{\bar{y}_1 - \bar{y}_2}{se}$) and calculate its value for the data set (call it t^*)
- ② Combine the $n_1 + n_2$ observations
- ③ Take a sample of size n_1 without replacement (to represent the Group C)
- ④ The remaining n_2 observations constitute the Group T
- ⑤ Compute the value of t (call it t_i) and repeat the process a large number (B) of times
- ⑥ **P-value:** $p = \Sigma I(t_{(i)} \geq t^*)/B$

THE PERMUTATION TEST



$$t_{(1)} < t_{(2)} < \dots < t_{(B)}$$

$$t^* = \frac{\bar{y}_1 - \bar{y}_2}{se}$$

$$\text{P-value: } p = \Sigma I(t_{(i)} \geq t^*)/B$$

THE BOOTSTRAP

- ➔ The Bootstrap tests are more widely applicable though less accurate than the permutation test
- ➔ Extremely useful for computing standard errors and confidence intervals

Suppose:

Experiment	
Trt 1	Trt 2
Y_{11}	Y_{21}
Y_{12}	Y_{22}
...	...
Y_{1n1}	Y_{2n2}
$\bar{Y}_1 \pm S_1$	$\bar{Y}_2 \pm S_2$

From distribution F

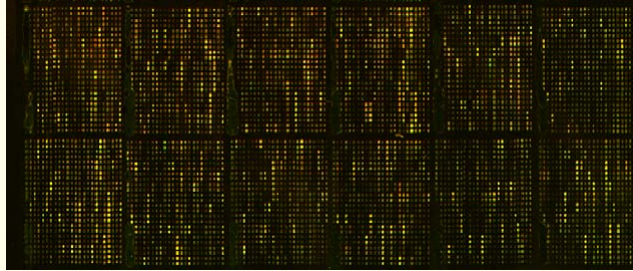
From distribution G

$H_0 : F = G$ vs. $H_1 : F \neq G$

THE BOOTSTRAP

- ① Define the statistics (e.g. $t = \frac{\bar{y}_1 - \bar{y}_2}{se}$) and calculate its value for the data set (call it t^*)
- ② Draw B samples of size $n_1 + n_2$ with replacement
- ③ The first n_1 observations constitute the Group C
- ④ The remaining n_2 observations constitute the Group T
- ⑤ Compute the value of t (call it t_i) and repeat the process a large number (B) of times
- ⑥ P-value: $p = \Sigma I(t_{(i)} \geq t^*)/B$

EXAMPLES OF TWO-SAMPLE MICROARRAY EXPERIMENTS

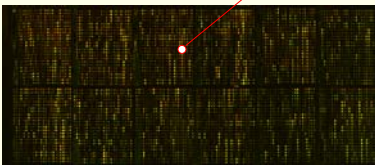


SINGLE-SLIDE MICROARRAY EXPERIMENTS

Two mRNA samples: Treatment (T) and Control (C)

{ Cy3 (G): mRNA Control sample
Cy5 (R): mRNA Treatment sample

(R,G) { R/G (expression ratio)
R x G (overall transcript abundance)



SINGLE-SLIDE METHODS

⇒ Schena et al. (1995):

- ① Normalization using spiked controls
- ② "Significance": $R/G \geq 5$ or $R/G \leq 0.2$ (fold change)

⇒ DeRisi et al. (1996):

- ① Normalization based on "housekeeping" genes
- ② "Significance" cutoff: ± 3 standard deviations (log scale)

⇒ Chen et al. (1997):

- ① Data dependent rule for choosing cutoffs for R/G
- ② Rule based on distributional assumptions for (R,G), including normality and constant CV.

SINGLE-SLIDE METHODS

⇒ Sapir and Churchill (2000):

- ① Background correction
- ② Normalization using robust regression of LogR vs LogG
- ③ Posterior probabilities of change using a mixture model for $\log(R/G)$

⇒ Newton et al. (2001):

- ① Hierarchical (Gamma-Gamma-Bernoulli) model for (R,G)
- ② Posterior odds of change as functions of (R+G) and RG (taking into account overall transcript abundance)

⇒ Hughes et al. (2000):

- ① Assumption that R and G $\overset{ind.}{\sim} N(\mu_i, \sigma_i^2)$; $\sigma_i^2 = f(\mu_i)$

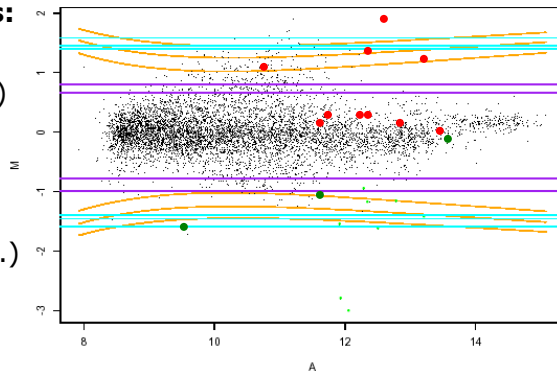
Single-slide Methods: Model dependent rule with cutoffs in the (LogR, LogG)-plane

MA-plot with contours:

Newton et al. (2001)
(odds: 1:1, 10:1, 100:1)

Chen et al. (1997)
(95% and 99% conf.)

Sapir, Churchill (2000)
(90, 95, 99% post. prob.)

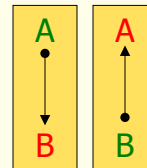


The points corresponding to genes adjusted p-value less than 0.05 (based on data from 16 slides) are colored in green (negative t-stat.) and red (positive t-stat.).

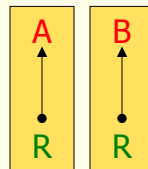
Adapted from Yang and Speed (2003)

TWO-SLIDE EXPERIMENTS

⇒ Dye-swap: mRNA samples A and B



⇒ Common reference: R



⇒ Notation:

- For each gene (spot), take $\log_2 A$ and $\log_2 B$ (and $\log_2 R$, if common reference)
- $\phi = \alpha - \beta$, where ϕ is the differential expression, $\alpha = E[\log_2 A]$, and $\beta = E[\log_2 B]$ across slides.
- $\sigma^2 = \text{Var}[\log_2(A/B)]$: constant variance across slides and samples (log ratio)

TWO-SLIDE EXPERIMENTS

⇒ Dye-swap: mRNA samples A and B

Direct estimate of ϕ :

$$\hat{\phi}_D = \frac{1}{2} \left[\log_2 \frac{A}{B} + \log_2 \frac{A'}{B'} \right] \quad \text{and} \quad \text{Var}(\hat{\phi}_D) = \frac{\sigma^2}{2}$$

⇒ Common reference: R

Indirect estimate of ϕ :

$$\hat{\phi}_I = \log_2 \frac{A}{R} + \log_2 \frac{B}{R} \quad \text{and} \quad \text{Var}(\hat{\phi}_I) = 2\sigma^2$$

REPLICATE SLIDE EXPERIMENTS

- n replicate hybridizations between samples A and B
- For each gene, we can compute \bar{M} and s^2 using the n log-ratios $M = \log_2(A/B)$

⇒ Classical approaches
(identifying differentially expressed genes)

- ① $|\bar{M}| \geq c$ ($c = \text{threshold}$)
 - Equal variability assumed across genes
- ② t-based statistics: $t = \sqrt{n} \bar{M} / s$; $|t|$ for two-sided tests
 - Gene-specific variances across slides

⇒ **Problems:** n quite small (outliers have large effect, s^2 may be very small by chance)

Suggestion: truncation of small std errors (e.g. 1%)

REPLICATE SLIDE METHODS

⇒ Efron et al. (2000):

something midway between a common and a gene specific standard error

$$t^* = \frac{\sqrt{n} \bar{M}}{a + s}$$

E.g.: 90th percentile of standard deviations

- Empirical Bayes approach

⇒ Lönnstedt and Speed (2001):

- ① Data from all genes are combined into estimates of a prior distribution
- ② These parameters are then combined at the gene level with mean and standard deviation to form a statistic B (Bayes log posterior odds)

REPLICATE SLIDE METHODS

⇒ Robust Models:

- Robust models to estimate ϕ , to attenuate the effect of outliers

⇒ Nonparametric Approaches:

- Tests based on ranks, for example (small sample sizes; discrete distribution of p-values)

⇒ Mixture Models:

- Background + foreground components (for R and G intensities)
- Differentially expressed + remainder genes distributions for log ratios
- Mixture distribution for p-values (Uniform + Beta)

REPLICATE SLIDE METHODS

⇒ Linear Models:

- ANOVA (fixed-effects) models (Kerr et al., 2000)
- Mixed-effects models (Wolfinger et al., 2001)

Models including terms for slide, dye, gene, treatment
(experimental group) and some interactions

$$\log(y_{ijk}) = \mu + A_i + D_j + V_k + G_g + (AG)_{ig} + (VG)_{kg} + \varepsilon_{ijk}$$

⇒ Error Models (Ideker et al., 2000):

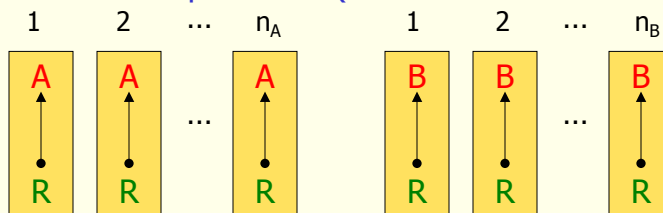
$$(VG_{kg} - VG_{k'g})$$

$$\begin{cases} T = \mu_T + \mu_T \varepsilon + \delta \\ C = \mu_C + \mu_C \varepsilon' + \delta' \end{cases} \Rightarrow \begin{pmatrix} \varepsilon \\ \varepsilon' \end{pmatrix} \sim N(0, E) \text{ and } \begin{pmatrix} \delta \\ \delta' \end{pmatrix} \sim N(0, D) \text{ across spots.}$$

- Likelihood ratio test for the hypothesis $\mu_T = \mu_C$ for each gene
- **Alternatives:** Roberts et al. (2000), Rocke and Durbin (2001), Theilhaber et al. (2001), Baggerly et al. (2001)

REPLICATE SLIDE METHODS

⇒ Indirect Comparisons (Common reference R):



⇒ *t*-statistic:

$$t = \frac{\bar{M}_A - \bar{M}_B}{s_p \sqrt{1/n_A + 1/n_B}}$$

pooled std deviation

- The same problems are found here, and the same suggestions are given:

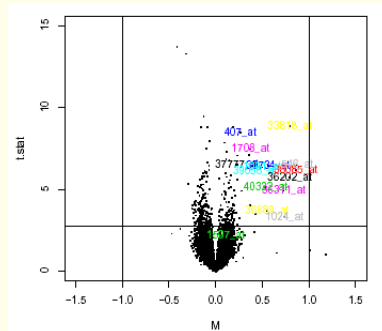
Modify *s_p* or use robust variants of $\bar{M}_A - \bar{M}_B$ and *t*.

Graphical Visualization of Results

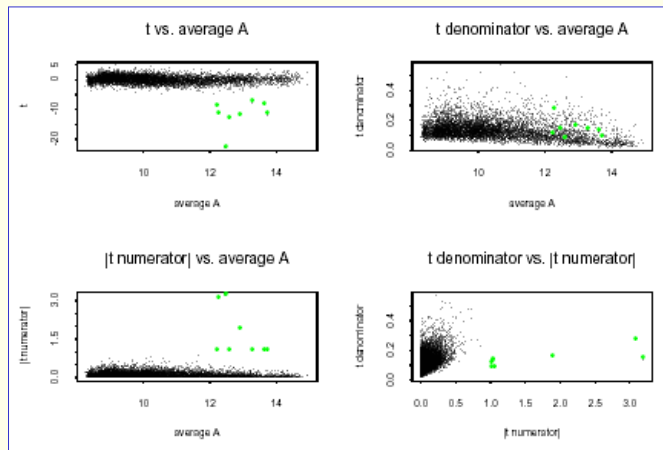
- Important features of the genes can be found examining graphical representations of the results, such as p-values, t-statistics, their numerators and denominators.
- Informal determination of cutoffs for declaring differentially expressed genes.
- Ranking genes by sets of statistics.

Example

Volcano Plot: $-\log(p\text{-values})$
against M (log ratio)



Example



Plots of t-statistics, numerator, and denominator, against overall intensity. The points corresponding to genes with adjusted p-value less than 0.01 are colored in green.

REFERENCES

- Efron, B., Storey, J. D., Tibshirani, R. (2001) Microarrays, empirical Bayes methods, and false discovery rates. Stanford Technical Report # 216.
- Efron, B., Tibshirani, R. (1998) An Introduction to the Bootstrap. Chapman & Hall.
- Kerr, M. K., Churchill, G. A. (2001) Statistical design and the analysis of gene expression microarray data. *Genet. Res. Camb.* 77: 123-128.
- Lee, M.-L. T., Kuo, F. C., Whitmore, G. A., Sklar, J. (2000) Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci. USA* 97: 9834-9839.
- Rosa, G. J. M. (1999) Robust mixed linear models in quantitative genetics: Bayesian analysis via Gibbs sampling. *Proc. International Symposium on Animal Breeding and Genetics*, Sept 21-24, p.133-160.
- Tusher, V. G., Tibshirani, R., Chu, G. (2001) Significance analysis of microarray applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* 98: 5116-5121.
- von Rohr, P., Hoeschele, I. (2002) Bayesian QTL mapping using skewed Student-t distributions. *Genetics, Selection, Evolution* 34 (1): 1-21.
- Tempelman, R. (1998) Generalized linear mixed models in dairy cattle breeding. *J. Dairy Sci.* 81: 1428-1444.
- Schena, M. et al. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270 (5235), 467-470.
- DeRisi, J. L. et al. (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genetics* 14, 457-460.
- Chen, Y., E. R. Dougherty and M. L. Bittner (1997) Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomedical Optics* 2, 364-374.
- Sapir, M. and G. A. Churchill (2000) Estimating the posterior probability of differential gene expression from microarray data (poster). The Jackson Laboratory (<http://www.jax.org/research/churchill/>).
- Newton, M. A. et al. (2000) On differential variability of expression ratios: improving statistical inference about gene expression changes from microarray data. *J. Computational Biol.* 8, 37-52.