Multivariate Statistical Methods in fMRI:

Thresholding:

The Devil is in the details

Daniel B. Rowe, PhD

Department of Biophysics
Medical College of Wisconsin
8701 Watertown Plank Rd.
Milwaukee, WI 53226
dbrowe@mcw.edu

Outline

+ **Dependencies**

+ **Distributions**

+ **One Voxel Statistics**

+ **One Voxel Thresholding**

+ **Two Voxel Statistics**

+ **Two Voxel Thresholding**

+ **Remarks**

Dependencies

Multivariate Bayesian Statistics, Rowe, D.B. CRC Press.

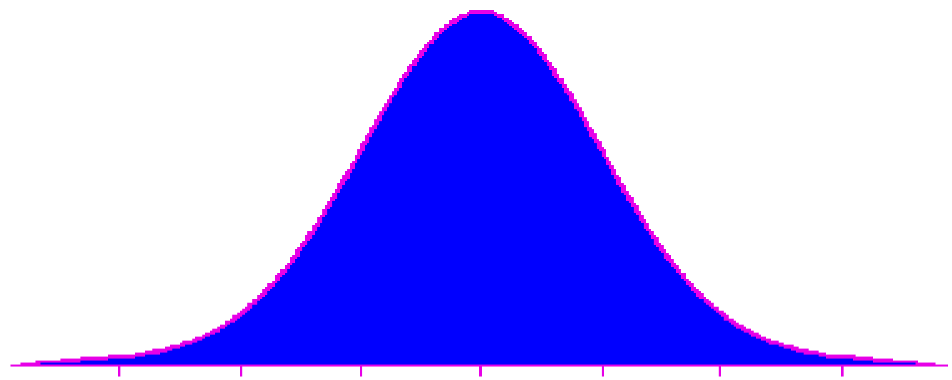
Independent is a special case of dependent.

We should safeguard against using an independent model and drawing inferences with it by explicitly modeling dependencies.

The routine analysis of fMRI data is to fit an independent multiple regression model in each voxel and compute a statistic of interest.

When computing "activation maps" these dependencies are apparent.

Normal Distribution

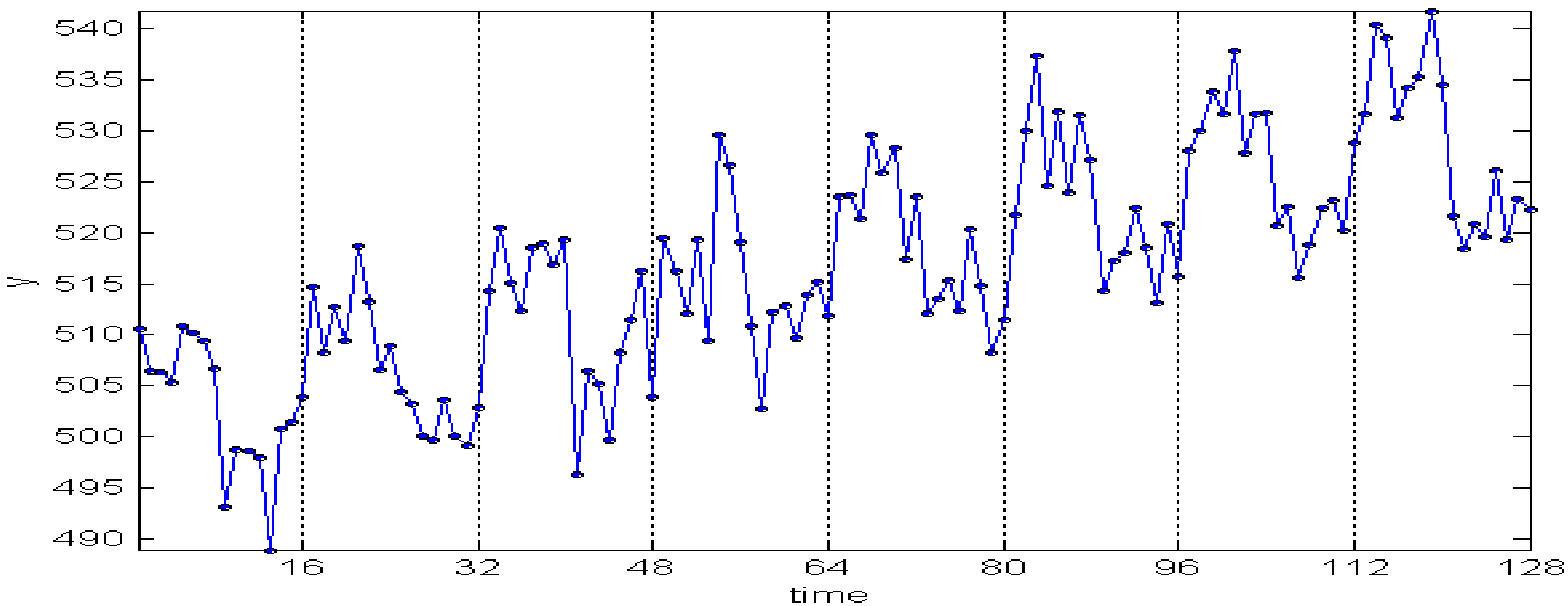


$$x : 1 \times 1, \qquad x \sim N(\mu, \sigma^2)$$

$$p(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Central Limit Theorem: The Normal distribution, also called the bell curve, is that distribution which any other distribution (with finite first and second moments) tends to be on average.

One Voxel: Block design finger tapping



$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_i, \quad \epsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$$
$$i = 1, \ldots, n$$

Normal Distributions

$$x : 1 \times 1, \qquad x \sim N(\mu, \sigma^2)$$

$$p(x|\mu, \sigma^2) \;=\; (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)}$$

Normal Distributions

$$x : 1 \times 1, \qquad x \sim N(\mu, \sigma^2)$$

$$p(x|\mu,\sigma^2) = (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)}$$

$$x : p \times 1, \qquad x \sim N(\mu, \Sigma)$$

$$p(x|\mu,\Sigma) = (2\pi)^{-\frac{p}{2}}|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$$

Normal Distributions

$$x : 1 \times 1, \qquad x \sim N(\mu, \sigma^2)$$

$$p(x|\mu, \sigma^2) = (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)}$$

$$x : p \times 1, \qquad x \sim N(\mu, \Sigma)$$

$$p(x|\mu, \Sigma) = (2\pi)^{-\frac{p}{2}}|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$$

$$X : p \times n, \qquad X \sim N(M, \Phi \otimes \Sigma)$$

$$p(X|M, \Sigma, \Phi) = (2\pi)^{-\frac{np}{2}}|\Phi|^{-\frac{p}{2}}|\Sigma|^{-\frac{n}{2}}e^{-\frac{1}{2}tr\Phi^{-1}(X-M)'\Sigma^{-1}(X-M)}$$

Normal Distributions

$$x : 1 \times 1, \qquad x \sim N(\mu, \sigma^2)$$

$$p(x|\mu, \sigma^2) = (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)}$$

$$x : p \times 1, \qquad x \sim N(\mu, \Sigma)$$

$$p(x|\mu, \Sigma) = (2\pi)^{-\frac{p}{2}}|\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$$

$$X : p \times n, \qquad X \sim N(M, \Phi \otimes \Sigma)$$

$$p(X|M, \Sigma, \Phi) = (2\pi)^{-\frac{np}{2}}|\Phi|^{-\frac{p}{2}}|\Sigma|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\Phi^{-1}(X-M)'\Sigma^{-1}(X-M)}$$

If $p = 1$ then second is scalar normal.
If $p = 1$, then third is vector normal.
If $n = 1$, then third is vector normal.
If $p = 1$ & $n = 1$, then third is scalar normal.

Normal Distributions

$$x : 1 \times 1, \qquad x \sim N(\mu, \sigma^2)$$

$$p(x|\mu, \sigma^2) = (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)}$$

$$x : p \times 1, \qquad x \sim N(\mu, \Sigma)$$

$$p(x|\mu, \Sigma) = (2\pi)^{-\frac{p}{2}}|\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$$

$$X : p \times n, \qquad X \sim N(M, \Phi \otimes \Sigma)$$

$$p(X|M, \Sigma, \Phi) = (2\pi)^{-\frac{np}{2}}|\Phi|^{-\frac{p}{2}}|\Sigma|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\Phi^{-1}(X-M)'\Sigma^{-1}(X-M)}$$

$$E(X|M, \Sigma, \Phi) = M$$

$$Mode(X|M, \Sigma, \Phi) = M$$

$$var(vec(X)|M, \Sigma, \Phi) = \Phi \otimes \Sigma$$

Student T-Distributions

$$t : 1 \times 1, \qquad t \sim t(\nu, t_0, \sigma^2, \phi^2)$$

$$p(t|\nu, t_0, \sigma^2, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} (\sigma^2)^{-1}}{[\phi^2 + \frac{1}{\nu}(t - t_0)(\sigma^2)^{-1}(t - t_0)]^{\frac{\nu+1}{2}}}$$

Student T-Distributions

$$t : 1 \times 1, \qquad t \sim t(\nu, t_0, \sigma^2, \phi^2)$$

$$p(t|\nu, t_0, \sigma^2, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} (\sigma^2)^{-1}}{\left[\phi^2 + \frac{1}{\nu}(t - t_0)(\sigma^2)^{-1}(t - t_0)\right]^{\frac{\nu+1}{2}}}$$

$$t : p \times 1, \qquad t \sim t(\nu, t_0, \Sigma, \phi^2)$$

$$p(t|\nu, t_0, \Sigma, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} |\Sigma|^{-\frac{1}{2}}}{\left[\phi^2 + \frac{1}{\nu}(t - t_0)'\Sigma^{-1}(t - t_0)\right]^{\frac{\nu+p}{2}}}$$

Student T-Distributions

$$t : 1 \times 1, \qquad t \sim t(\nu, t_0, \sigma^2, \phi^2)$$

$$p(t|\nu, t_0, \sigma^2, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} (\sigma^2)^{-1}}{[\phi^2 + \frac{1}{\nu}(t - t_0)(\sigma^2)^{-1}(t - t_0)]^{\frac{\nu+1}{2}}}$$

$$t : p \times 1, \qquad t \sim t(\nu, t_0, \Sigma, \phi^2)$$

$$p(t|\nu, t_0, \Sigma, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} |\Sigma|^{-\frac{1}{2}}}{[\phi^2 + \frac{1}{\nu}(t - t_0)'\Sigma^{-1}(t - t_0)]^{\frac{\nu+p}{2}}}$$

$$T : p \times n, \qquad T \sim T(\nu, T_0, \Sigma, \Phi)$$

$$p(T|\nu, T_0, \Sigma, \Phi) = k_T \frac{|\Phi|^{\frac{\nu}{2}} |\Sigma|^{-\frac{n}{2}}}{|\Phi + \frac{1}{\nu}(T - T_0)'\Sigma^{-1}(T - T_0)|^{\frac{\nu+p}{2}}}$$

Student T-Distributions

$$t : 1 \times 1, \qquad t \sim t(\nu, t_0, \sigma^2, \phi^2)$$

$$p(t|\nu, t_0, \sigma^2, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}}(\sigma^2)^{-1}}{[\phi^2 + \frac{1}{\nu}(t - t_0)(\sigma^2)^{-1}(t - t_0)]^{\frac{\nu+1}{2}}}$$

$$t : p \times 1, \qquad t \sim t(\nu, t_0, \Sigma, \phi^2)$$

$$p(t|\nu, t_0, \Sigma, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}}|\Sigma|^{-\frac{1}{2}}}{[\phi^2 + \frac{1}{\nu}(t - t_0)'\Sigma^{-1}(t - t_0)]^{\frac{\nu+p}{2}}}$$

$$T : p \times n, \qquad T \sim T(\nu, T_0, \Sigma, \Phi)$$

$$p(T|\nu, T_0, \Sigma, \Phi) = k_T \frac{|\Phi|^{\frac{\nu}{2}}|\Sigma|^{-\frac{n}{2}}}{|\Phi + \frac{1}{\nu}(T - T_0)'\Sigma^{-1}(T - T_0)|^{\frac{\nu+p}{2}}}$$

If $p = 1$ then second is scalar t.

If $p = 1$, then third is vector t.

If $n = 1$, then third is vector t.

If $p = 1$ & $n = 1$, then third is scalar t.

Student T-Distributions

$$t : 1 \times 1, \qquad t \sim t(\nu, t_0, \sigma^2, \phi^2)$$

$$p(t|\nu, t_0, \sigma^2, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} (\sigma^2)^{-1}}{\left[\phi^2 + \frac{1}{\nu}(t - t_0)(\sigma^2)^{-1}(t - t_0)\right]^{\frac{\nu+1}{2}}}$$

$$t : p \times 1, \qquad t \sim t(\nu, t_0, \Sigma, \phi^2)$$

$$p(t|\nu, t_0, \Sigma, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} |\Sigma|^{-\frac{1}{2}}}{\left[\phi^2 + \frac{1}{\nu}(t - t_0)'\Sigma^{-1}(t - t_0)\right]^{\frac{\nu+p}{2}}}$$

$$T : p \times n, \qquad T \sim T(\nu, T_0, \Sigma, \Phi)$$

$$p(T|\nu, T_0, \Sigma, \Phi) = k_T \frac{|\Phi|^{\frac{\nu}{2}} |\Sigma|^{-\frac{n}{2}}}{\left|\Phi + \frac{1}{\nu}(T - T_0)'\Sigma^{-1}(T - T_0)\right|^{\frac{\nu+p}{2}}}$$

$$E(T|\nu, T_0, \Sigma, \Phi) = T_0,$$

$$Mode(T|\nu, T_0, \Sigma, \Phi) = T_0,$$

$$var(vec(T)|\nu, T_0, \Sigma, \Phi) = \frac{\nu}{\nu - 2}(\Phi \otimes \Sigma)$$

Wishart (Gamma) Distributions

$$g : 1 \times 1, \qquad g \sim \Gamma(\alpha, \beta)$$

$$p(g|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\,\beta^{\alpha}} g^{\alpha-1} e^{-g/\beta}$$

Wishart Distributions

$$g : 1 \times 1, \qquad g \sim \Gamma(\alpha, \beta)$$

$$p(g|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\,\beta^{\alpha}} g^{\alpha-1} e^{-g/\beta}, \quad \alpha = \frac{\nu_0}{2}, \; \beta = 2\upsilon^2$$

$$g : 1 \times 1, \qquad g \sim W(\upsilon^2, 1, \nu_0)$$

$$p(g|\upsilon^2, \nu_0) = k_W (\upsilon^2)^{-\frac{\nu_0}{2}} g^{\frac{\nu_0 - 2}{2}} e^{-\frac{1}{2}(\upsilon^2)^{-1} g}$$

$$k_W^{-1} = \Gamma\left(\frac{\nu_0}{2}\right) 2^{\frac{\nu_0}{2}}$$

Wishart Distributions

$$g : 1 \times 1, \qquad g \sim \Gamma(\alpha, \beta)$$

$$p(g|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\,\beta^{\alpha}} g^{\alpha-1} e^{-g/\beta}, \quad \alpha = \frac{\nu_0}{2}, \;\; \beta = 2\upsilon^2$$

$$g : 1 \times 1, \qquad g \sim W(\upsilon^2, 1, \nu_0)$$

$$p(g|\upsilon^2, \nu_0) = k_W (\upsilon^2)^{-\frac{\nu_0}{2}} g^{\frac{\nu_0-2}{2}} e^{-\frac{1}{2}(\upsilon^2)^{-1} g}$$

$$G : p \times p, \qquad G \sim W(\Upsilon, p, \nu_0)$$

$$p(G|\Upsilon, p, \nu_0) = k_W |\Upsilon|^{-\frac{\nu_0}{2}} |G|^{\frac{\nu_0-p-1}{2}} e^{-\frac{1}{2} tr \Upsilon^{-1} G}$$

$$k_W^{-1} = \Gamma\left(\frac{\nu_0}{2}\right) 2^{\frac{\nu_0}{2}}$$

$$k_W^{-1} = 2^{\frac{\nu_0 p}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma\left(\frac{\nu_0 + 1 - j}{2}\right)$$

If $p = 1$ then the Wishart reduces to the Gamma.

Wishart Distributions

$$g : 1 \times 1, \qquad g \sim \Gamma(\alpha, \beta)$$

$$p(g|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} g^{\alpha-1} e^{-g/\beta}, \quad \alpha = \frac{\nu_0}{2}, \ \beta = 2\upsilon^2$$

$$g : 1 \times 1, \qquad g \sim W(\upsilon^2, 1, \nu_0)$$

$$p(g|\upsilon^2, \nu_0) = k_W (\upsilon^2)^{-\frac{\nu_0}{2}} g^{\frac{\nu_0 - 2}{2}} e^{-\frac{1}{2}(\upsilon^2)^{-1} g}$$

$$G : p \times p, \qquad G \sim W(\Upsilon, p, \nu_0)$$

$$p(G|\Upsilon, p, \nu_0) = k_W |\Upsilon|^{-\frac{\nu_0}{2}} |G|^{\frac{\nu_0 - p - 1}{2}} e^{-\frac{1}{2} tr \Upsilon^{-1} G}$$

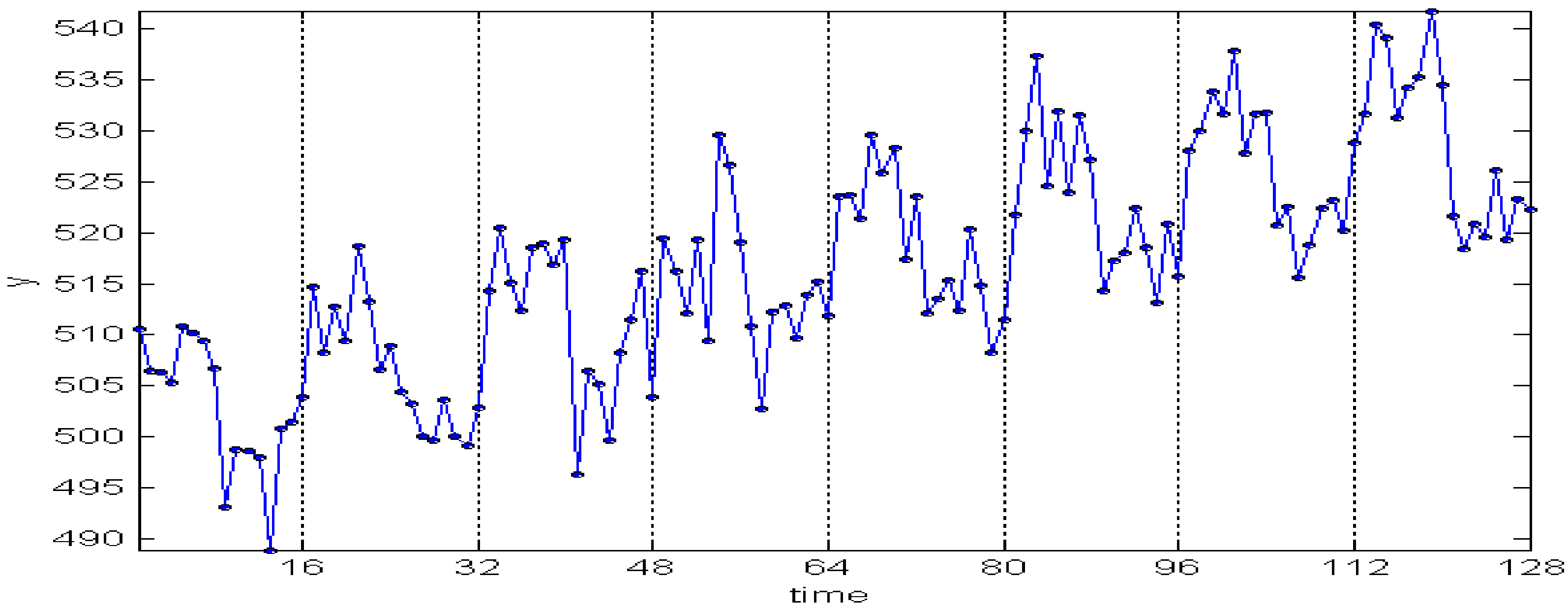
$$E(G|\nu_0, \Upsilon) = \nu_0 \Upsilon$$

$$Mode(G|\nu_0, \Upsilon) = (\nu_0 - p - 1)\Upsilon$$

$$var(g_{ij}|\nu_0, \Upsilon) = \nu_0(\upsilon_{ij}^2 + \upsilon_{ii}\upsilon_{jj})$$

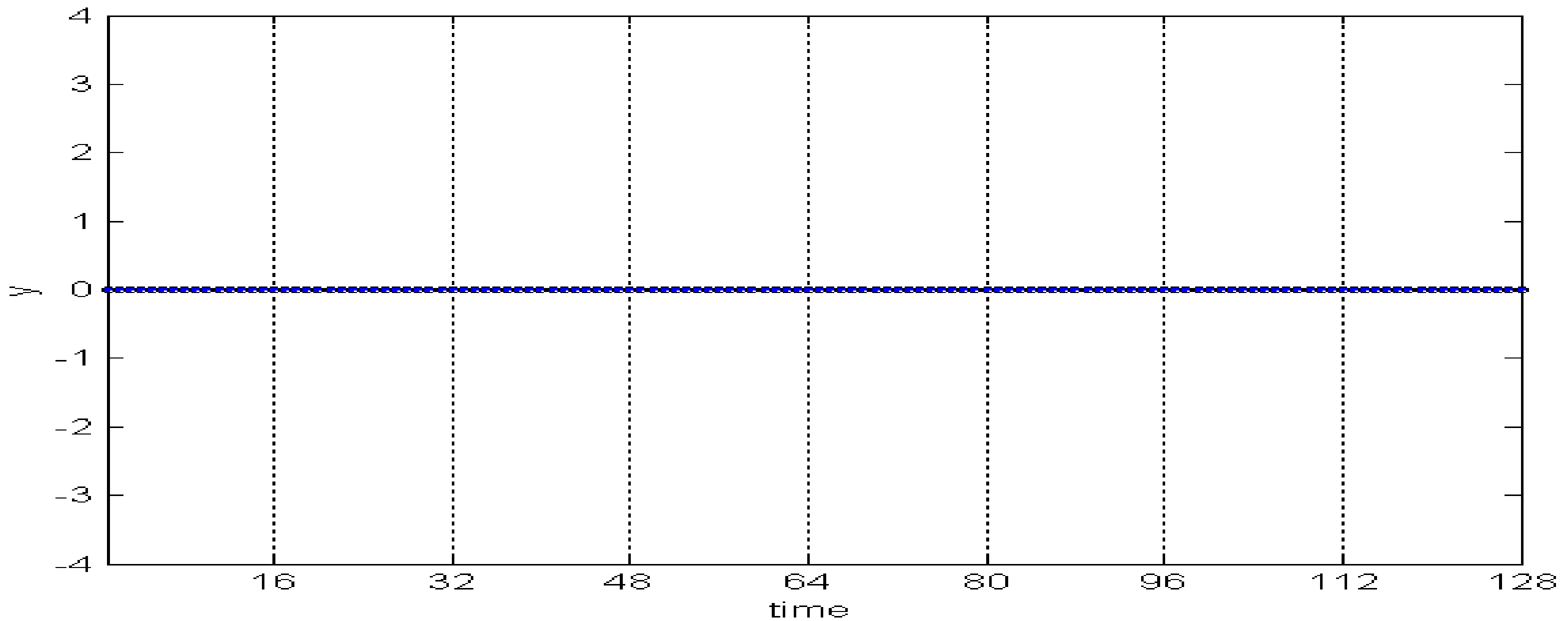
$$cov(g_{ij}g_{kl}|\nu_0, \Upsilon) = \nu_0(\upsilon_{ik}\upsilon_{jl} + \upsilon_{il}\upsilon_{jk})$$

One Voxel: Block design finger tapping



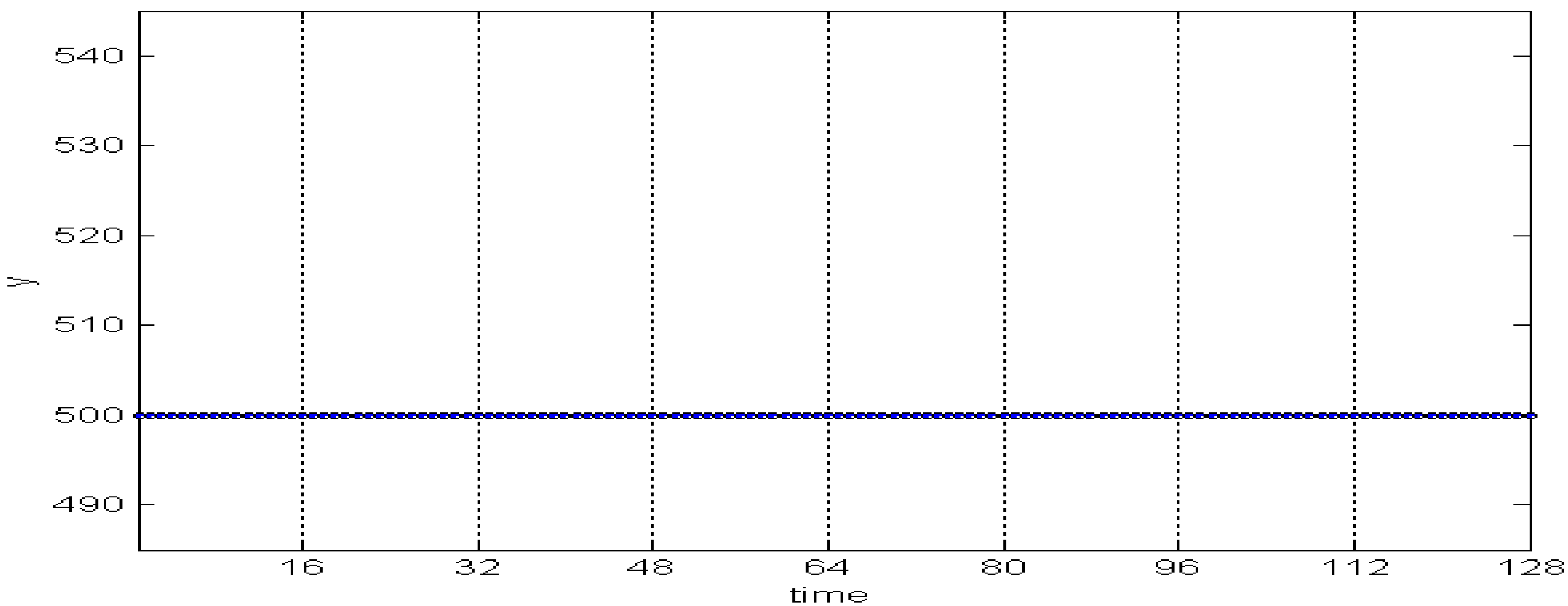
$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_i, \quad \epsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$$
$$i = 1, \ldots, n$$

One Voxel Cont.



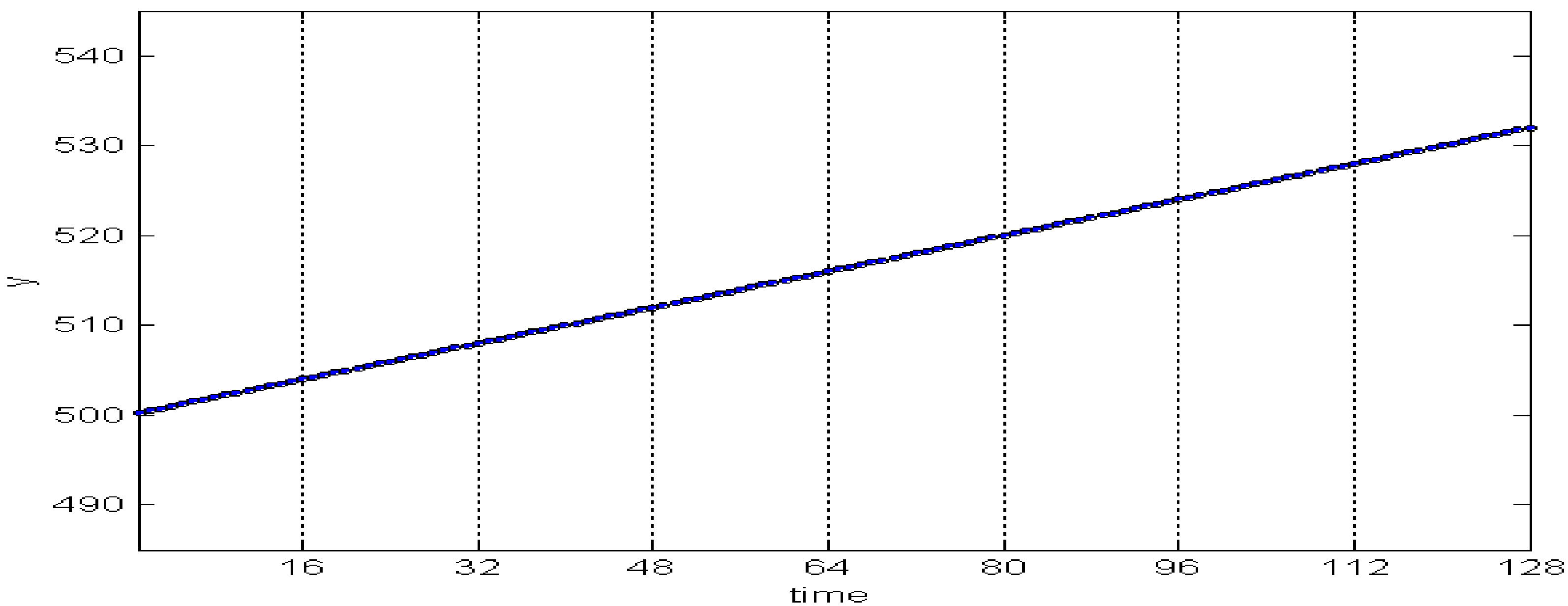
$$y_i = 0$$
$$i = 1, \ldots, n$$

One Voxel Cont.



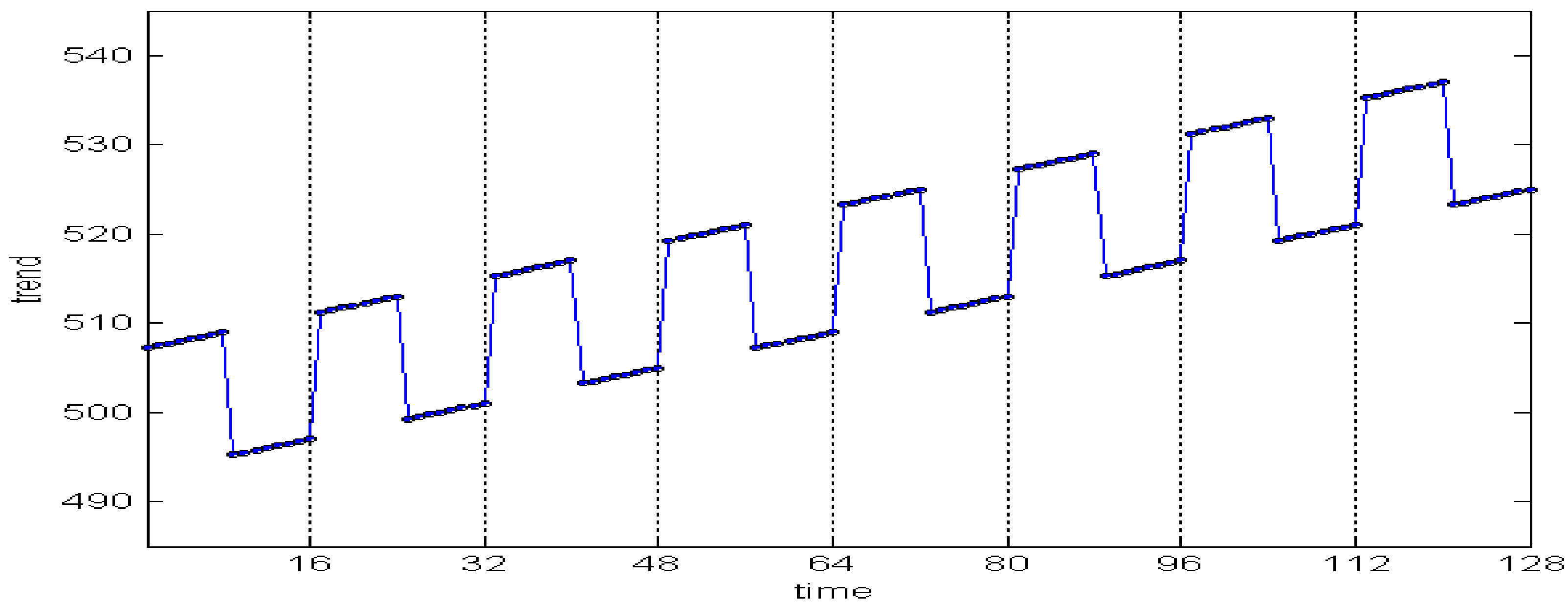
$$y_i = 500$$
$$i = 1, \ldots, n$$

One Voxel Cont.



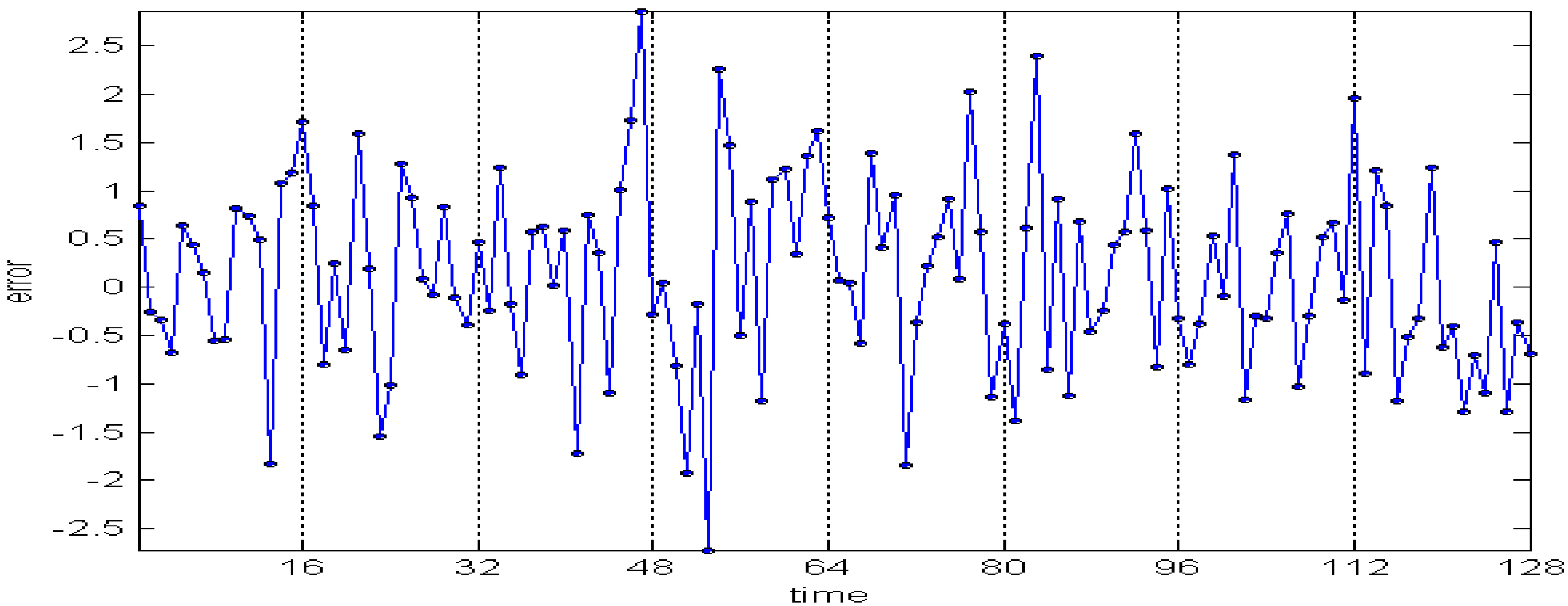
$$y_i = 500 + .25(i)$$
$$i = 1, \ldots, n$$

One Voxel Cont.



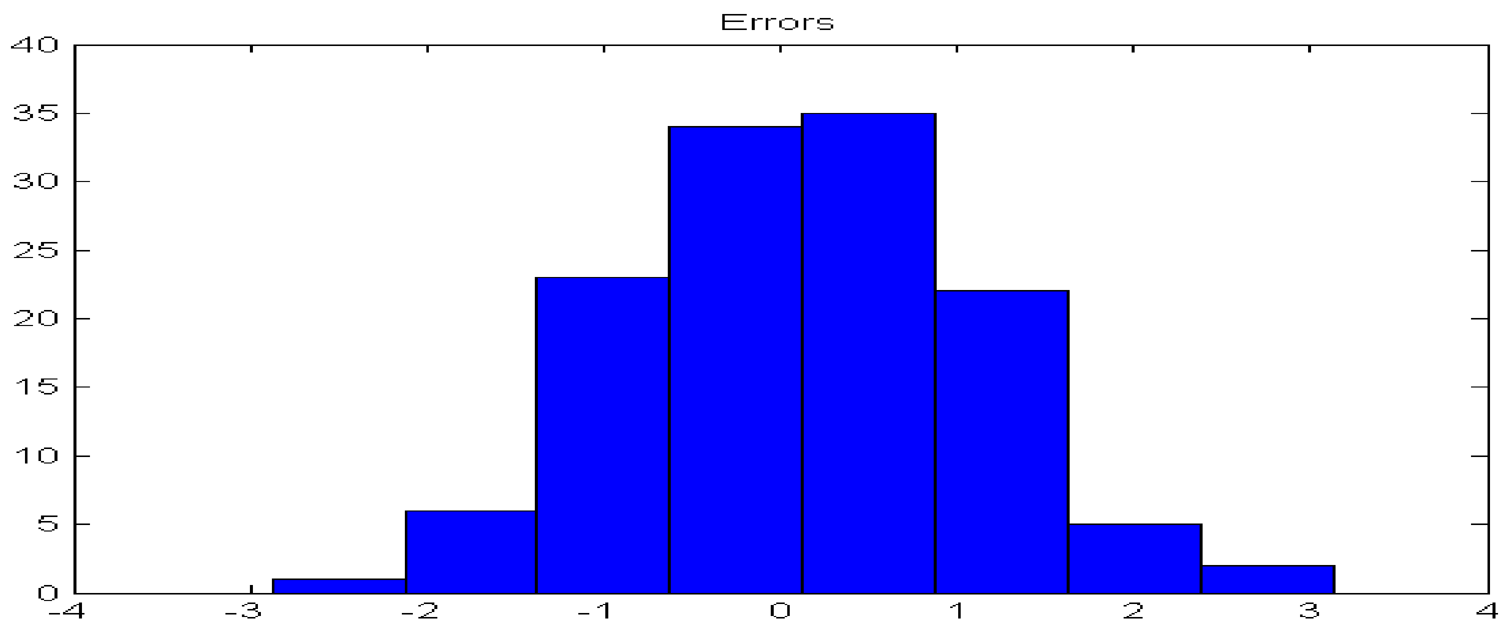
$$y_i = 500 + .25(i) + 7x_{2i}$$
$$i = 1, \ldots, n$$
$$x_{2i} = -1 \text{ if rest}, 1 \text{ if task.}$$

One Voxel Cont.



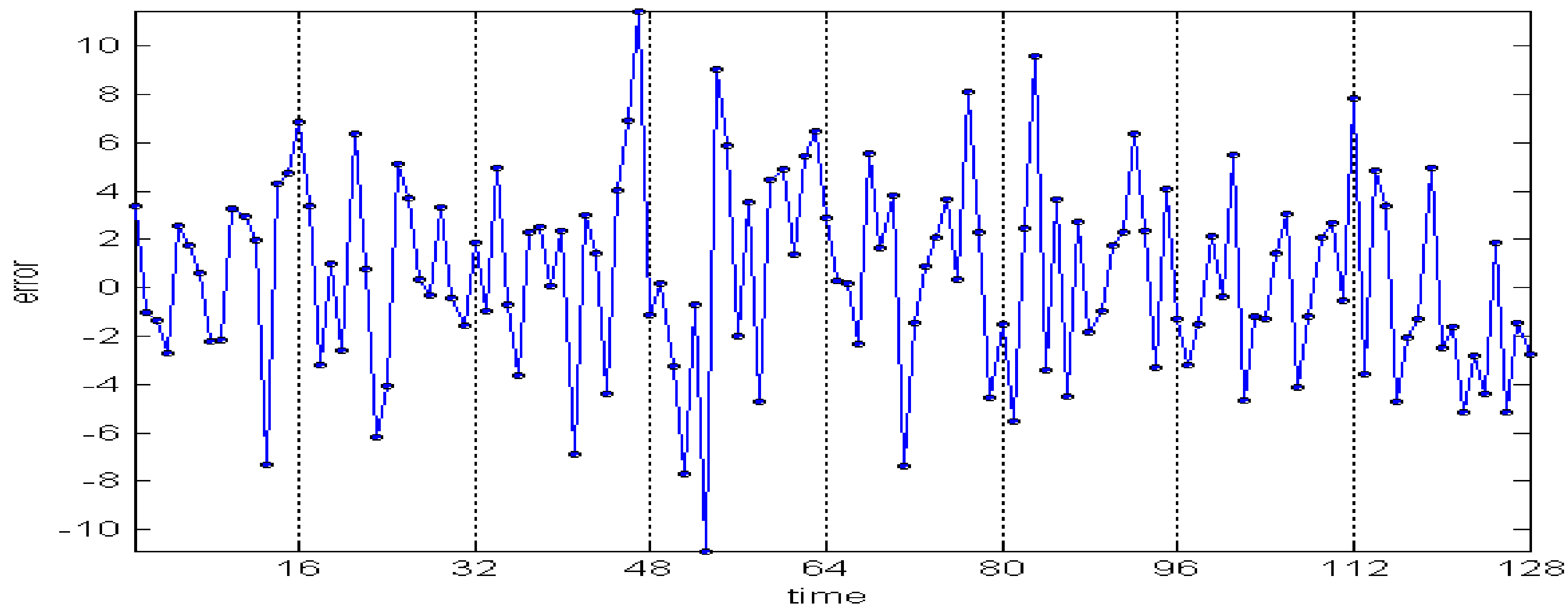
$$\delta_i \overset{\text{iid}}{\sim} N(0,1)$$
$$i = 1, \ldots, n$$

One Voxel Cont.



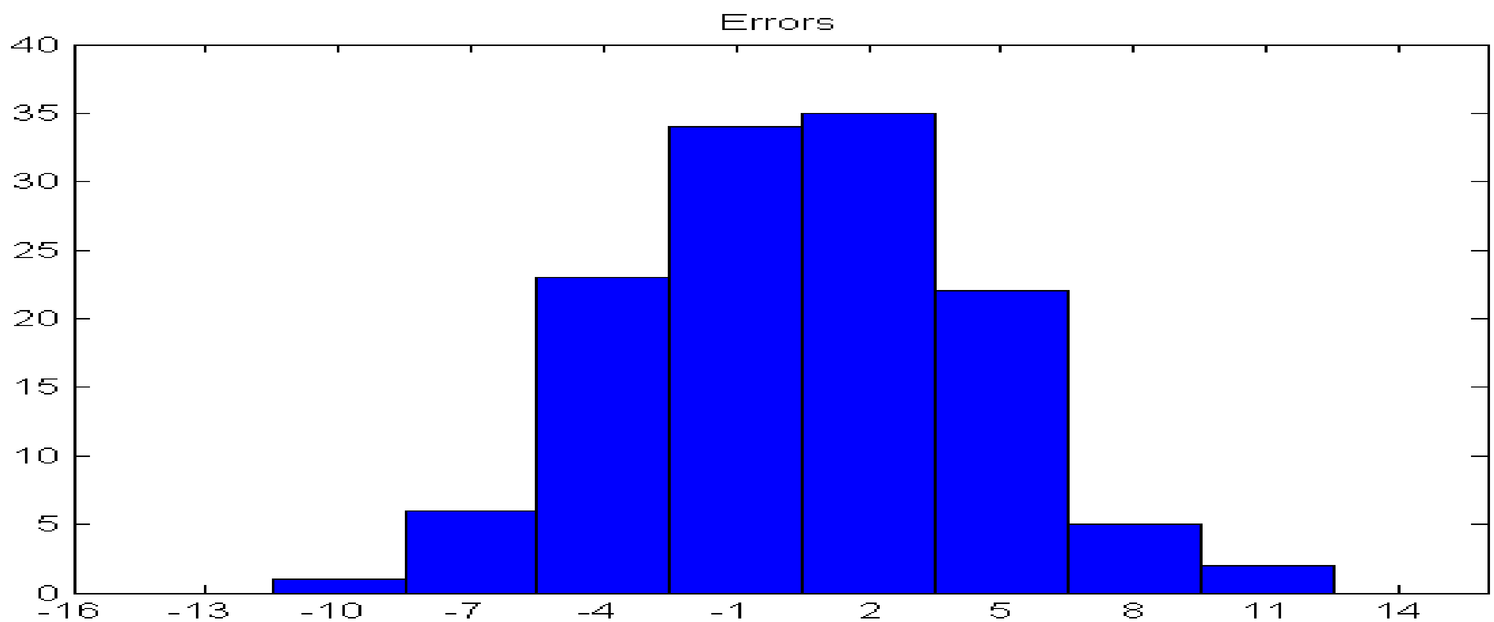
δ_i Histogram

One Voxel Cont.



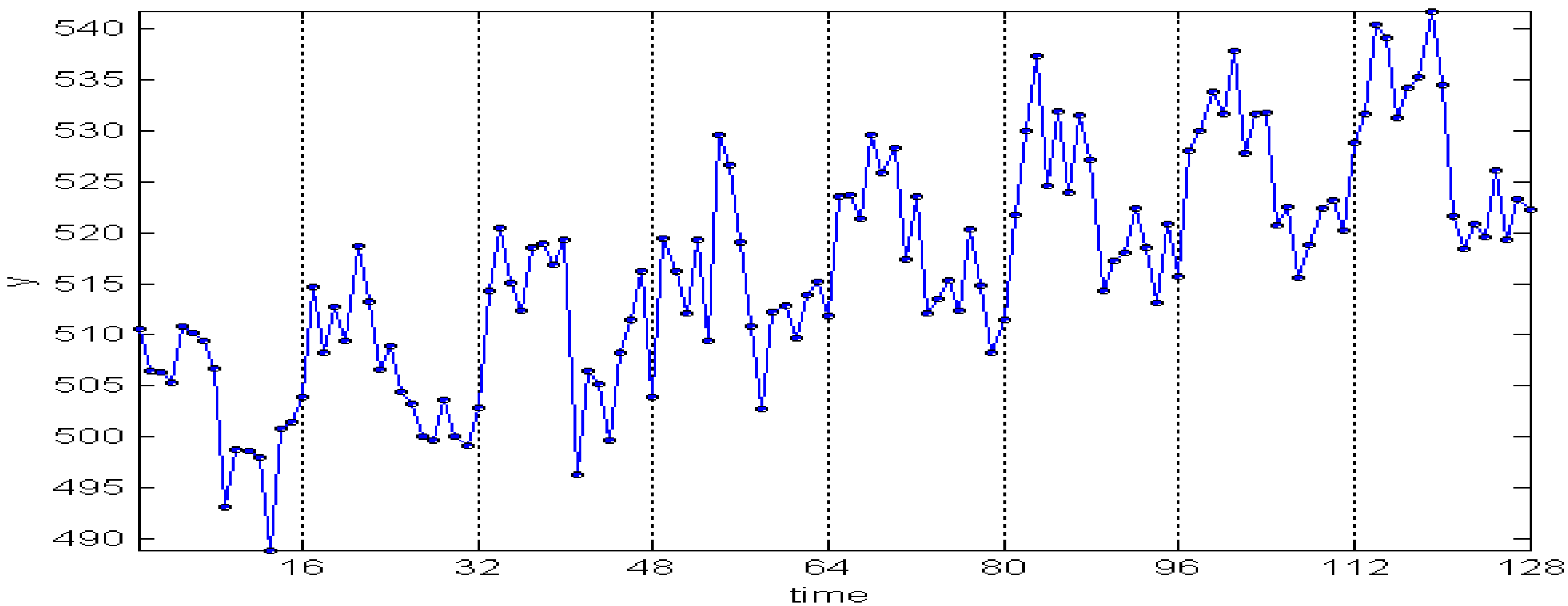
$$\epsilon_i = 4\delta_i, \quad \epsilon_i \overset{\text{iid}}{\sim} N(0, 4^2)$$
$$i = 1, \ldots, n$$

One Voxel Cont.



ϵ_i Histogram

One Voxel Cont.



$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_i, \quad \epsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$$
$$i = 1, \ldots, n$$

One Voxel Cont.

$$\underset{n\times 1}{y} = \underset{n\times 3}{X}\ \underset{3\times 1}{\alpha} + \underset{n\times 1}{\epsilon}, \qquad \underset{n\times 1}{\epsilon} \sim N(0,\sigma^2 I_n)$$

Where

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad x_i = \begin{pmatrix} 1 \\ x_{1i} \\ x_{2i} \end{pmatrix}, \quad X = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$$

$$\alpha = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$
\begin{array}{|c|}\hline \text{obs} \\ \hline y_1 \\ \vdots \\ y_8 \\ y_9 \\ \vdots \\ y_{16} \\ \vdots \\ \vdots \\ y_{113} \\ \vdots \\ y_{120} \\ y_{121} \\ \vdots \\ y_{128} \\ \hline \end{array}
=
\begin{array}{|rrr|}\hline \text{con} & \text{lin} & \text{ref} \\ \hline 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 8 & 1 \\ 1 & 9 & -1 \\ \vdots & \vdots & \vdots \\ 1 & 16 & -1 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & 113 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 120 & 1 \\ 1 & 121 & -1 \\ \vdots & \vdots & \vdots \\ 1 & 128 & -1 \\ \hline \end{array}
*
\begin{array}{|c|}\hline \text{coef} \\ \hline \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \hline \end{array}
+
\begin{array}{|c|}\hline \text{error} \\ \hline \epsilon_1 \\ \vdots \\ \epsilon_8 \\ \epsilon_9 \\ \vdots \\ \epsilon_{16} \\ \vdots \\ \vdots \\ \epsilon_{113} \\ \vdots \\ \epsilon_{120} \\ \epsilon_{121} \\ \vdots \\ \epsilon_{128} \\ \hline \end{array}
$$

One Voxel Cont.

$$\underset{n\times 1}{y} = \underset{n\times 3}{X}\ \underset{3\times 1}{\alpha} + \underset{n\times 1}{\epsilon}, \qquad \underset{n\times 1}{\epsilon} \sim N(0,\sigma^2 I_n)$$

$$\begin{aligned} p(y|\alpha,\sigma^2,X) &= (2\pi)^{-\frac{n}{2}}|\sigma^2 I_n|^{-\frac{1}{2}} e^{-\frac{1}{2}(y-X\alpha)'(\sigma^2 I_n)^{-1}(y-X\alpha)} \\ &= (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} e^{-\frac{(y-X\alpha)'(y-X\alpha)}{2\sigma^2}} \end{aligned}$$

It can be shown that

$$\underset{3\times 1}{\hat{\alpha}} = (X'X)^{-1}X'y$$

and

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y-X\hat{\alpha})'(y-X\hat{\alpha}) = \frac{1}{n}g$$

are MLE's.

One Voxel Cont.

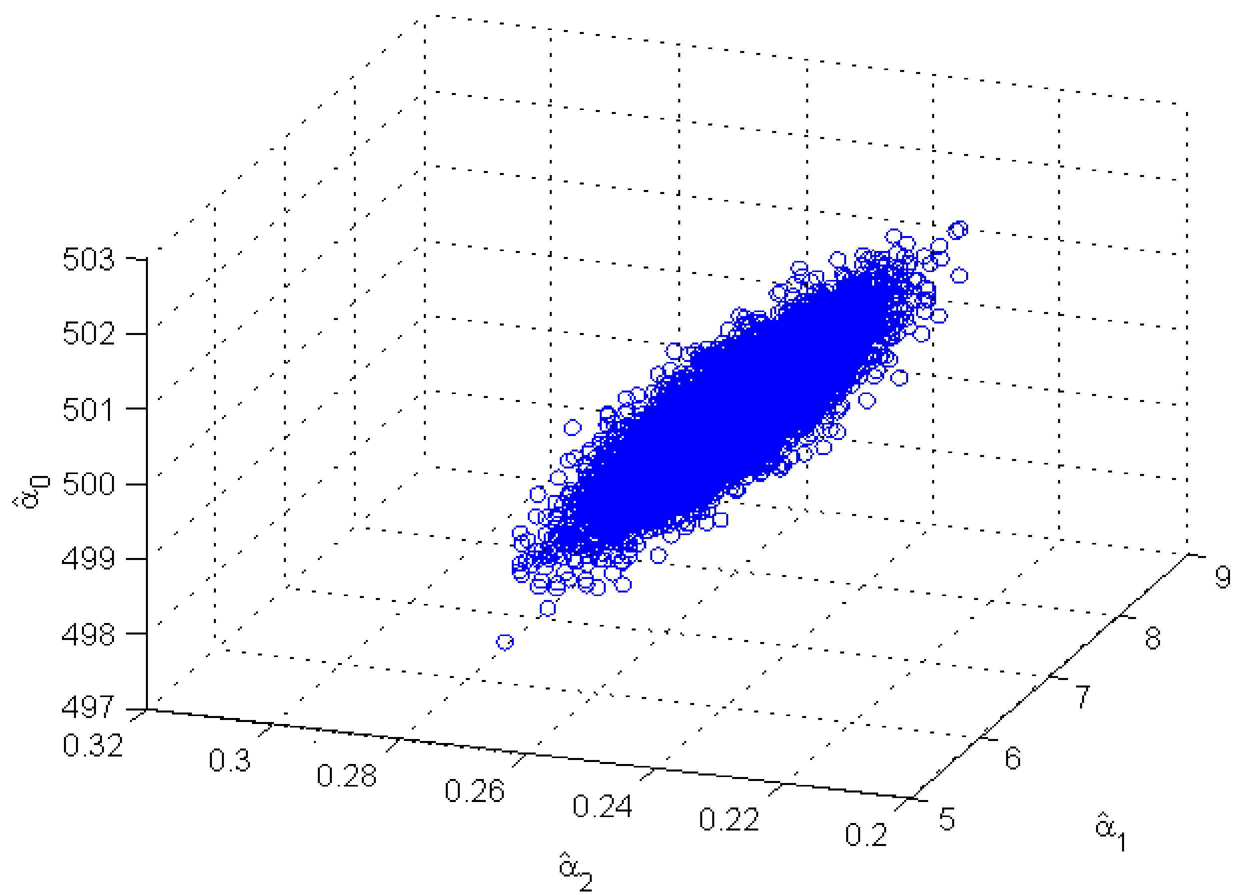
The values for $\hat{\alpha}$ and $\hat{\sigma}^2$ for the previous plotted data are.

$$
\begin{aligned}
\hat{\alpha} &= (501.3976, 0.2368, 6.2765)' \\
g &= 1934.0 \\
\hat{\sigma}_n^2 &= 15.1090 \\
\tilde{\sigma}_{n-3}^2 &= 15.4716
\end{aligned}
$$

One Voxel Cont.

Repeat steps to get the previous one voxel data 10000 times.

Scatter plot of the $\hat{\alpha}$'s. Not the same each time.



Trivariate distribution.

One Voxel Cont.

It can be shown that the distribution of $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)'$

$$\underset{3\times 1}{\hat{\alpha}} \sim t(n-2-1, \alpha, \tilde{\sigma}^2(X'X)^{-1}, 1)$$

$$p(\hat{\alpha}) = \frac{k_t \left[\tilde{\sigma}^2(X'X)^{-1}\right]^{-\frac{1}{2}}}{\left\{1 + \frac{1}{n-3}(\hat{\alpha}-\alpha)' \left[\tilde{\sigma}^2(X'X)^{-1}\right]^{-1}(\hat{\alpha}-\alpha)\right\}^{\frac{(n-3)+3}{2}}}$$

where

$$\tilde{\sigma}^2 = \frac{1}{n-3}(y - X\hat{\alpha})'(y - X\hat{\alpha}).$$

e_n is n-dim vec of 1's, c_n is n-dim vec. of count. $\#$'s, r_n is ref funct.

$$X'X = \begin{pmatrix} e_n'e_n & e_n'c_n & e_n'r_n \\ c_n'e_n & c_n'c_n & c_n'r_n \\ r_n'e_n & r_n'c_n & r_n'r_n \end{pmatrix} = \begin{pmatrix} 128 & 8256 & 0 \\ 8256 & 707264 & -512 \\ 0 & -512 & 128 \end{pmatrix}$$

One Voxel Cont.

Compare to the previous multivariate t-distribution

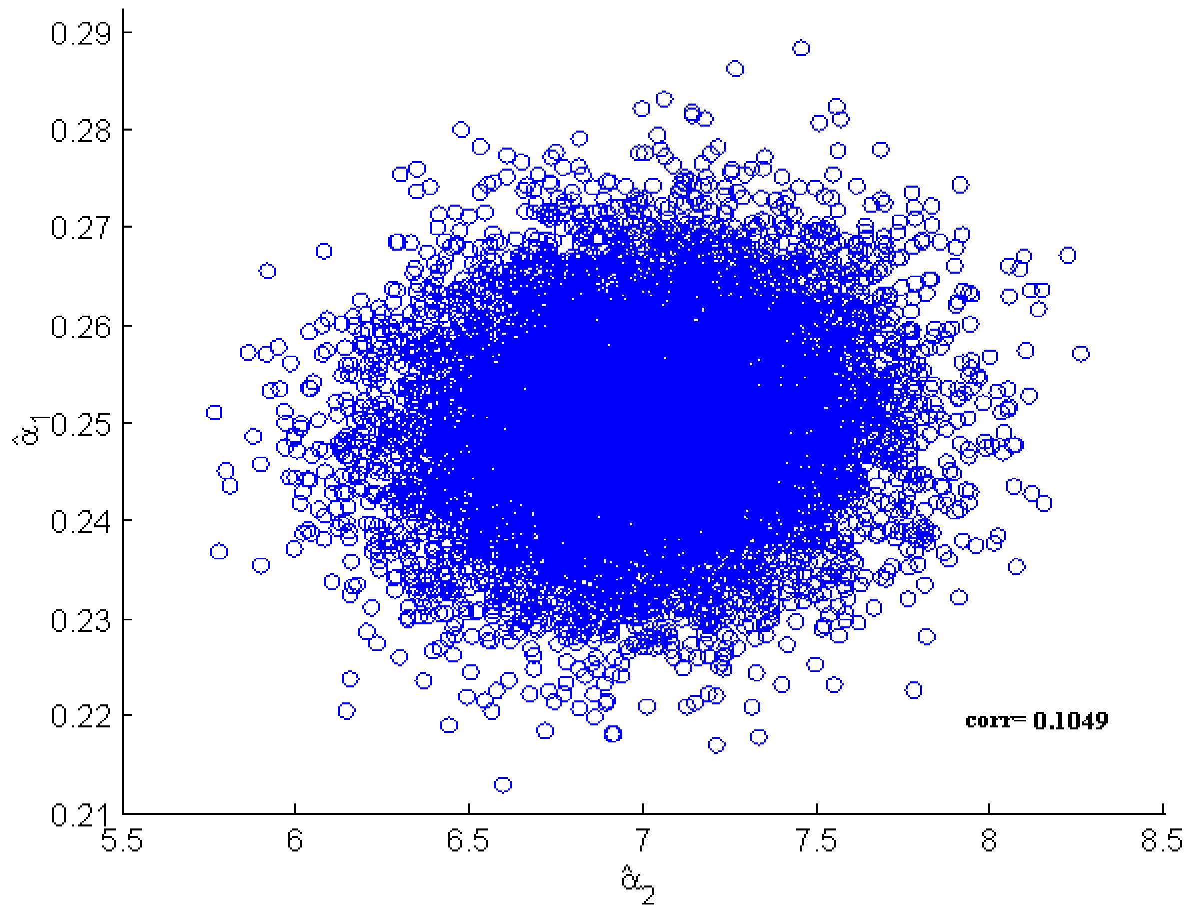
$$\underset{p \times 1}{t} \sim t(\nu, t_0, \Sigma, \phi^2)$$

$$p(t|\nu, t_0, \Sigma, \phi^2) = \frac{k_t(\phi^2)^{\frac{\nu}{2}}|\Sigma|^{-\frac{1}{2}}}{[\phi^2 + \frac{1}{\nu}(t - t_0)'\Sigma^{-1}(t - t_0)]^{\frac{\nu+p}{2}}}$$

One Voxel Cont.

Scatter plot of the $\hat{\alpha}_1$ and $\hat{\alpha}_2$'s.

Gravity in the $\hat{\alpha}_0$ direction.



Bivariate distribution. Histogram.

One Voxel Cont.

We can compute the marginal distribution of $\hat{\alpha}_* = (\hat{\alpha}_1, \hat{\alpha}_2)'$

$$p(\hat{\alpha}_1, \hat{\alpha}_2) = \int p(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2) \; d\hat{\alpha}_0.$$

It can be shown that

$$\underset{2 \times 1}{(\hat{\alpha}_1, \hat{\alpha}_2)'} \sim t(n-2-1, \alpha_*, \tilde{\sigma}^2 W_{**}, 1)$$

$$p(\hat{\alpha}_*) = \frac{k_t \left[\tilde{\sigma}^2 W_{**}\right]^{-\frac{1}{2}}}{\left\{1 + \frac{1}{n-3} (\hat{\alpha}_* - \alpha_*)' \left[\tilde{\sigma}^2 W_{**}\right]^{-1} (\hat{\alpha}_* - \alpha_*)\right\}^{\frac{(n-3)+2}{2}}}$$

where

$$W = (X'X)^{-1} \qquad E(\hat{\alpha}_*) = \alpha_*$$

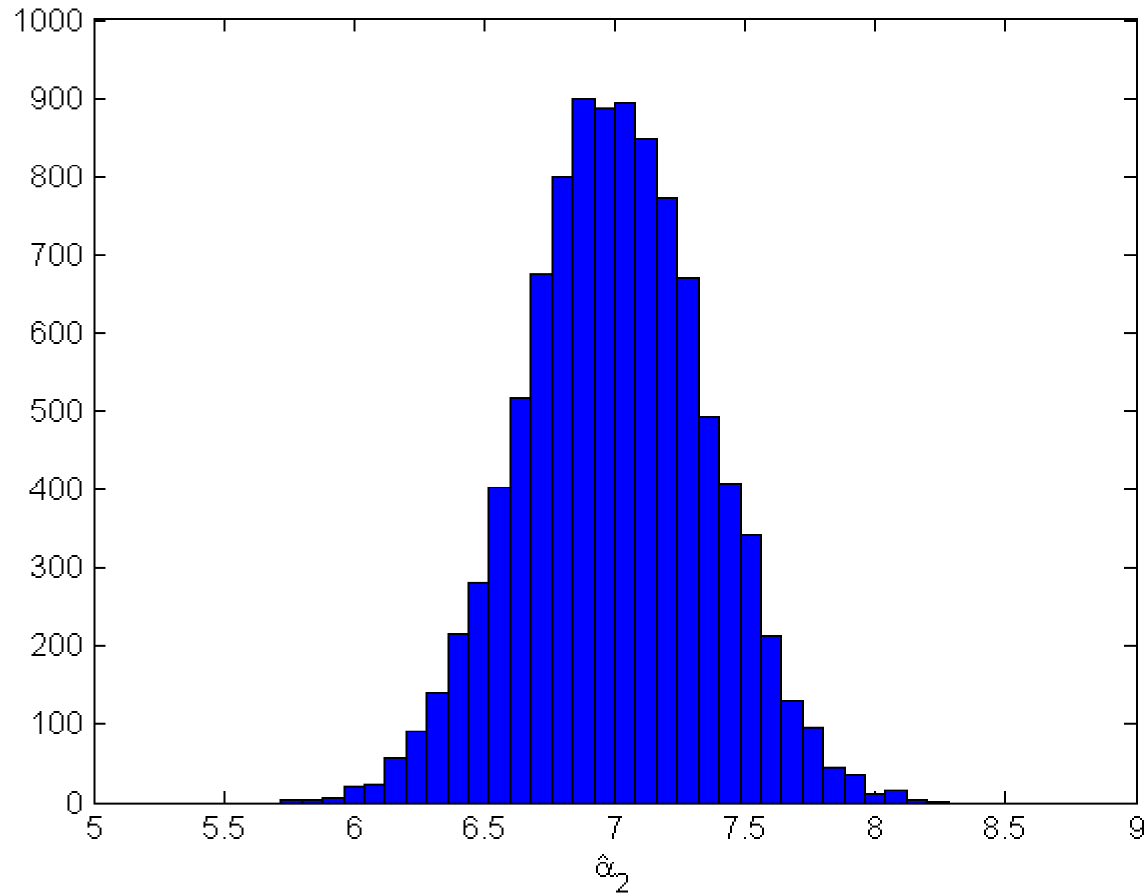
$$= \begin{pmatrix} W_{11} & W_{1*} \\ W_{*1} & W_{**} \end{pmatrix} \qquad var(\hat{\alpha}_*) = \frac{(n-2-1)}{(n-2-1)-2} \tilde{\sigma}^2 W_{**}$$

W_{**} is the lower right 2×2 matrix.

One Voxel Cont.

Histogram of the $\hat{\alpha}_2$'s.

Gravity in the $\hat{\alpha}_1$ direction.



Univariate distribution.

	mean	variance
Hist	7.0008	0.1229
True	7	0.1285

One Voxel Cont.

We can compute the marginal distribution of $\hat{\alpha}_2$

$$p(\hat{\alpha}_2) = \int\int p(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)\ d\hat{\alpha}_0\ d\hat{\alpha}_1.$$

It can be shown that

$$\underset{1\times 1}{\hat{\alpha}_2} \sim t(n-2-1, \alpha, \tilde{\sigma}^2 W_{33}, 1)$$

$$p(\hat{\alpha}_2) = \frac{k_t \left[\tilde{\sigma}^2 W_{33}\right]^{-\frac{1}{2}}}{\left\{1 + \frac{1}{n-3}(\hat{\alpha}_2 - \alpha_2)\left[\tilde{\sigma}^2 W_{33}\right]^{-1}(\hat{\alpha}_2 - \alpha_2)\right\}^{\frac{(n-3)+1}{2}}}$$

where $W = (X'X)^{-1}$ and W_{33} is the 33 element.

One Voxel Cont.

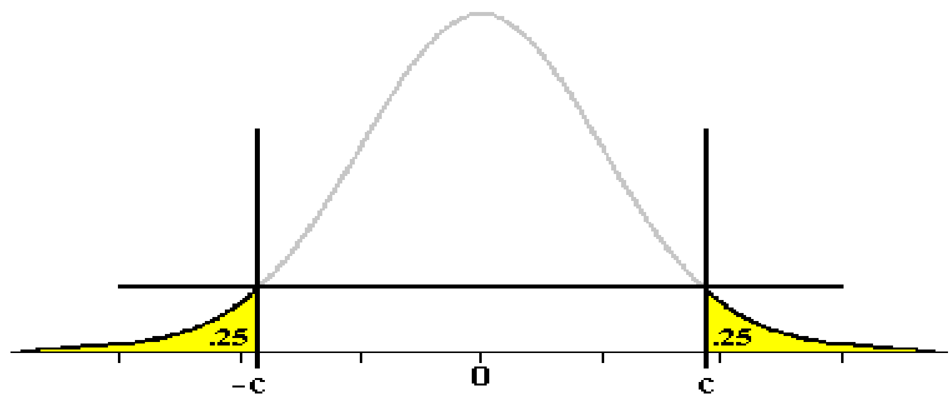
Compare to the previous student t-distribution

$$\underset{1\times 1}{t} \sim t(\nu, t_0, \sigma^2, \phi^2)$$

$$p(t|\nu, t_0, \sigma^2, \phi^2) = \frac{k_t(\phi^2)^{\frac{\nu}{2}}(\sigma^2)^{-\frac{1}{2}}}{\left[\phi^2 + \frac{1}{\nu}(t-t_0)(\sigma^2)^{-1}(t-t_0)\right]^{\frac{\nu+1}{2}}}$$

One Voxel Thresholding

Lower a parallel line and get a $(1-\alpha)*100\%$ CI.

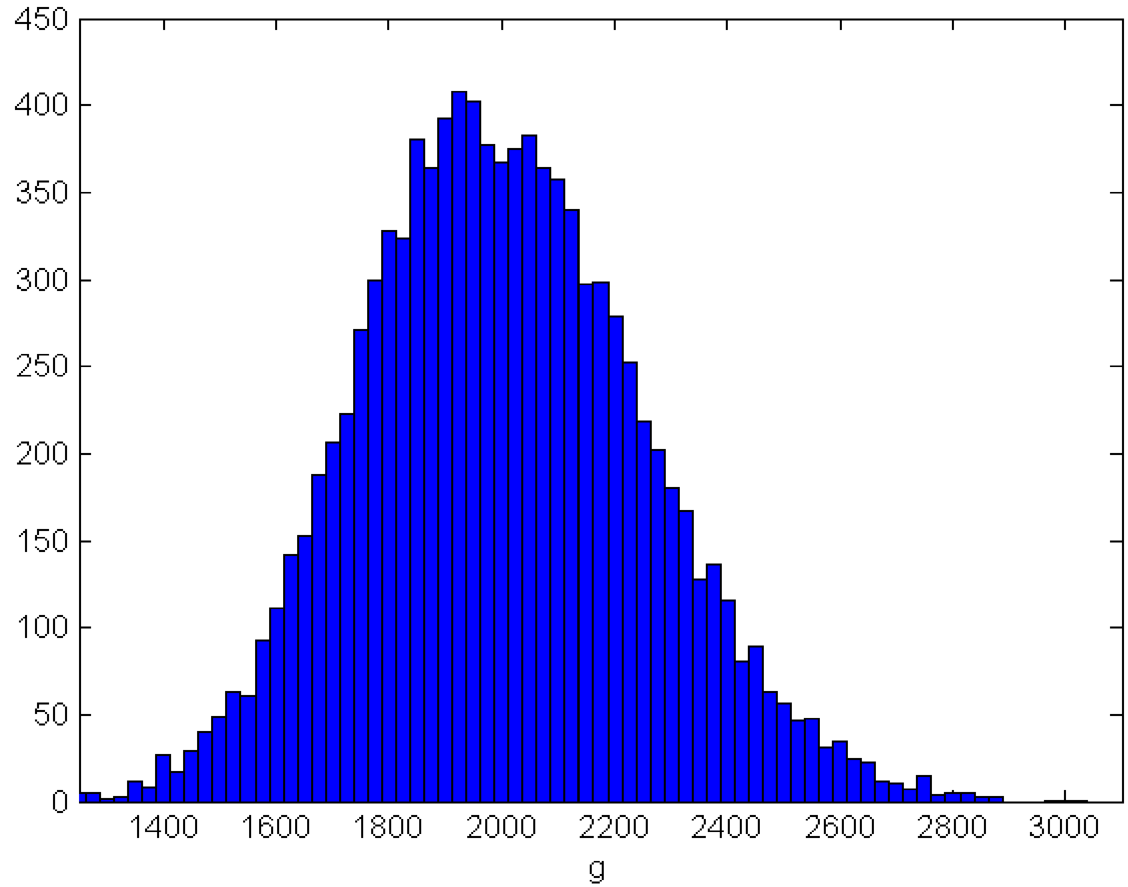


Hypothesis test of $\hat{\alpha}_2$ (Thresholding).

$$\hat{t} = \frac{\hat{\alpha}_2 - \alpha_2}{\sqrt{\tilde{\sigma}^2 W_{33}}} = \frac{6.2765 - 0}{\sqrt{(15.4716)(0.0079)}} = 17.9472$$

Look at t-distribution $t(n-2-1, 0, 1, 1)$ and if the test statistic is in the upper or lower tail with a certain probability reject H_0. $P(|t| > c) = \alpha$. Note that $F = t^2$.

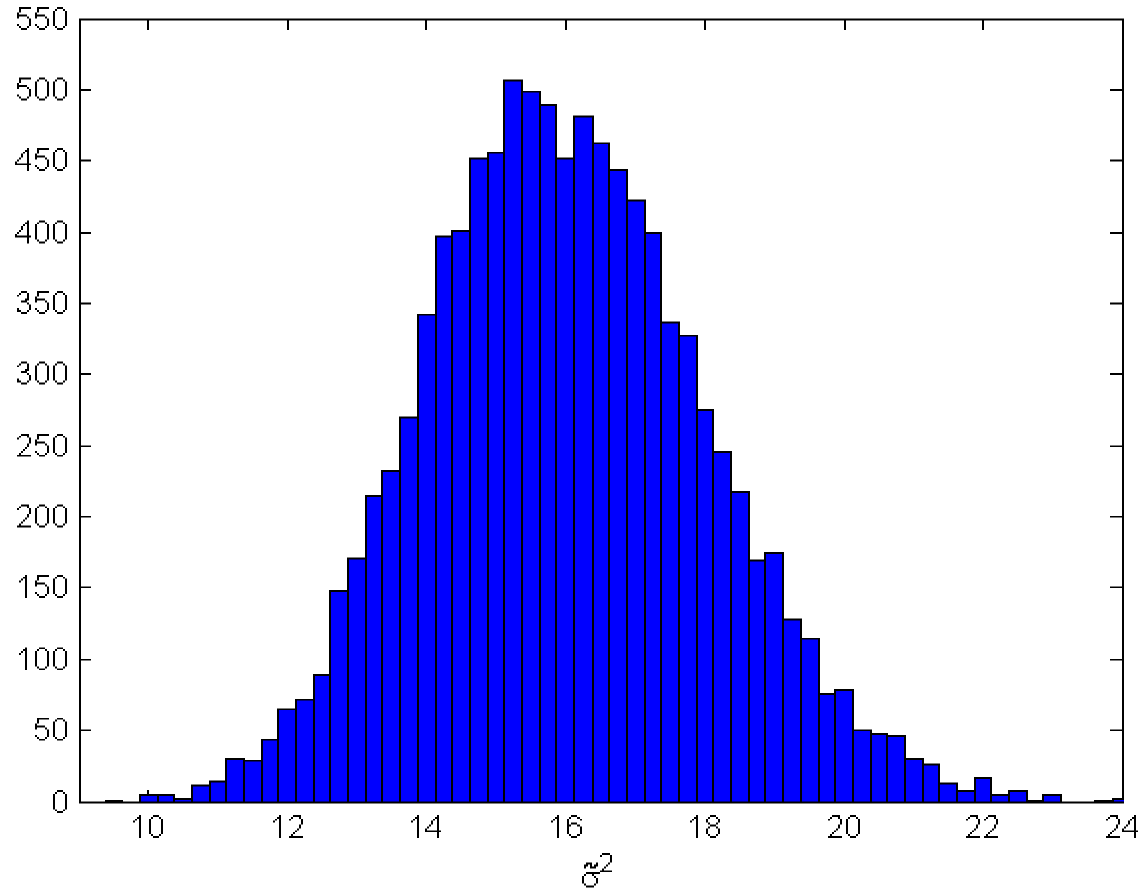
One Voxel Cont.



Histogram of the g's.

	mean	variance
Hist	2000.13	63049.29
True	2000	64000

One Voxel Cont.



Histogram of the $\tilde{\sigma}^2$'s.

	mean	variance
Hist	16.0011	4.0352
True	16	4

One Voxel Cont.

It can be shown that the distribution of $g = (y - X\hat{\alpha})'(y - X\hat{\alpha})$ is a scalar Wishart (Gamma)

$$\underset{1\times 1}{g} \sim W(\sigma^2, 1, n-2-1)$$

$$p(g) = k_W(\sigma^2)^{-\frac{(n-2-1)}{2}} g^{\frac{(n-2-1)-2}{2}} e^{-\frac{1}{2}(\sigma^2)^{-1} g}$$

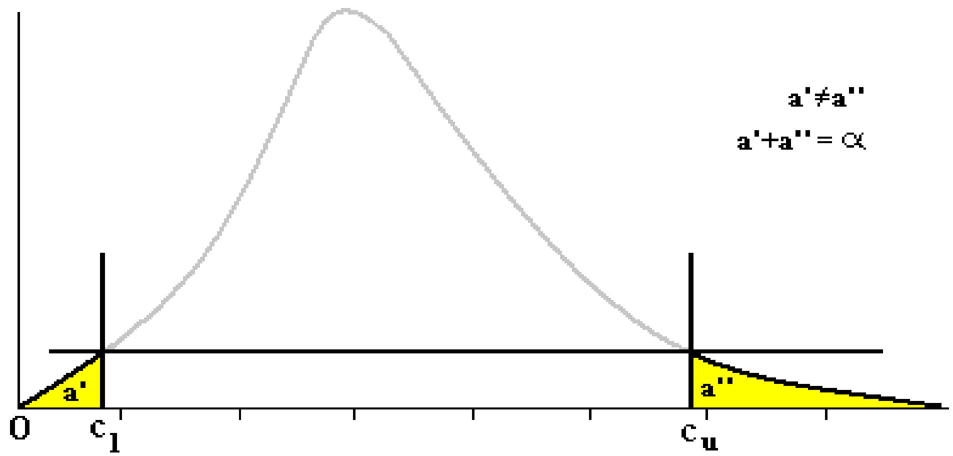
Compare to

$$g : 1 \times 1, \qquad g \sim W(\upsilon^2, 1, \nu_0)$$

$$p(g|\upsilon^2, \nu_0) = k_W(\upsilon^2)^{-\frac{\nu_0}{2}} g^{\frac{\nu_0 - 2}{2}} e^{-\frac{1}{2}(\upsilon^2)^{-1} g}$$

One Voxel Cont.

Lower a parallel line and get a $(1-\alpha)*100\%$ CI.



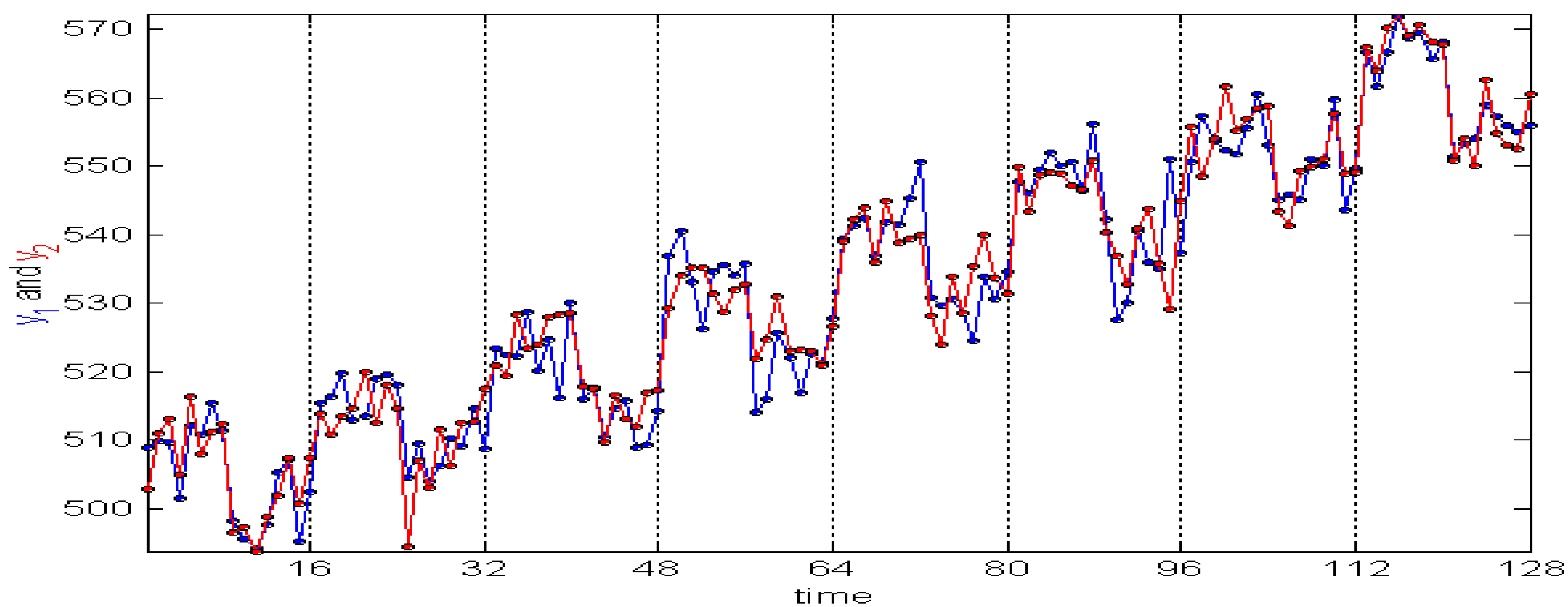
Hypothesis test of $\tilde{\sigma}^2$ or g.

$$\chi^2 = \frac{(n-2-1)\tilde{\sigma}^2}{\sigma^2} = \frac{(125)(15.4716)}{10} = 193.3950$$

Look at scalar Wishart (Gamma) distribution $W(10, 1, n-2-1)$ and if the test statistic is in the upper or lower tail with a certain probability reject H_0.

Minimal length CI's using Lagrange multipliers (Tate & Klett, JASA).

Two Voxels: Block design finger tapping



$$
\begin{aligned}
y_{1i} &= \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_{1i}, \\
y_{2i} &= \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \epsilon_{2i}, \\
i &= 1, \ldots, n
\end{aligned}
$$

Two Voxel Cont.

$$\underset{2\times 1}{y_i} = \underset{2\times 3}{x_i'} \; \underset{3\times 1}{B} + \underset{2\times 1}{\epsilon_i}, \qquad \underset{2\times 1}{\epsilon_i} \sim N(0,\Sigma)$$

Where

$$y_i = \begin{pmatrix} y_{1i} \\ y_{2i} \end{pmatrix}, \quad x_i = \begin{pmatrix} 1 \\ x_{1i} \\ x_{2i} \end{pmatrix}, \quad B = (\alpha, \gamma), \quad \epsilon_i = \begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \end{pmatrix}$$

$$\alpha = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}, \quad \gamma = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix}$$

Two Voxel Cont.

If we stack the y_i's

$$\underset{2n \times 1}{y} = \underset{2n \times 6}{(I_2 \otimes X)} \underset{6 \times 1}{\beta} + \underset{2n \times 1}{\epsilon}, \qquad \underset{2n \times 1}{\epsilon} \sim N(0, I_n \otimes \Sigma)$$

Where

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}, \quad \beta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$p(y|X, \beta, \Sigma) = (2\pi)^{-\frac{np}{2}} |I_n \otimes \Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}[y-(I_2 \otimes X)\beta]'(I_n \otimes \Sigma)^{-1}[y-(I_2 \otimes X)\beta]}$$

$\otimes$ is the Kronecker product which multiplies every element of its first matrix argument by its entire second matrix argument.

Two Voxel Cont.

It can be shown that the previous likelihood simplifies to

$$p(Y|X,B,\Sigma)=(2\pi)^{-\frac{np}{2}}|I_n|^{-\frac{p}{2}}|\Sigma|^{-\frac{n}{2}}e^{-\frac{1}{2}tr I_n^{-1}(Y-XB)'\Sigma^{-1}(Y-XB)}$$

Where

$$Y=\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix}, \quad X=\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}, \quad B=(\alpha,\gamma), \quad E=\begin{pmatrix} \epsilon_1' \\ \vdots \\ \epsilon_n' \end{pmatrix}$$

$$\underset{n\times 2}{Y} = \underset{n\times 3}{X}\ \underset{3\times 2}{B} + \underset{n\times 2}{E}, \qquad \underset{n\times 2}{E} \sim N(0, I_n \otimes \Sigma)$$

obs1	obs2
$y_{1,1}$	$y_{1,1}$
$\vdots$	$\vdots$
$y_{1,8}$	$y_{1,8}$
$y_{1,9}$	$y_{1,9}$
$\vdots$	$\vdots$
$y_{1,16}$	$y_{1,16}$
$\vdots$	$\vdots$
$\vdots$	$\vdots$
$y_{1,113}$	$y_{1,113}$
$\vdots$	$\vdots$
$y_{1,120}$	$y_{1,120}$
$y_{1,121}$	$y_{1,121}$
$\vdots$	$\vdots$
$y_{1,128}$	$y_{1,128}$

$=$

con	lin	ref
1	1	1
$\vdots$	$\vdots$	$\vdots$
1	8	1
1	9	-1
$\vdots$	$\vdots$	$\vdots$
1	16	-1
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$
1	113	1
$\vdots$	$\vdots$	$\vdots$
1	120	1
1	121	-1
$\vdots$	$\vdots$	$\vdots$
1	128	-1

$*$

cf1	cf2
α_0	γ_0
α_1	γ_1
α_2	γ_2

$+$

er1	er2
$\epsilon_{1,1}$	$\epsilon_{2,1}$
$\vdots$	$\vdots$
$\epsilon_{1,8}$	$\epsilon_{2,8}$
$\epsilon_{1,9}$	$\epsilon_{2,9}$
$\vdots$	$\vdots$
$\epsilon_{1,16}$	$\epsilon_{2,16}$
$\vdots$	$\vdots$
$\vdots$	$\vdots$
$\epsilon_{1,113}$	$\epsilon_{2,113}$
$\vdots$	$\vdots$
$\epsilon_{1,120}$	$\epsilon_{2,120}$
$\epsilon_{1,121}$	$\epsilon_{2,121}$
$\vdots$	$\vdots$
$\epsilon_{1,128}$	$\epsilon_{2,128}$

Two Voxels Cont.

With the likelihood (distribution)

$$p(Y|X,B,\Sigma)=(2\pi)^{-\frac{np}{2}}|I_n|^{-\frac{p}{2}}|\Sigma|^{-\frac{n}{2}}e^{-\frac{1}{2}tr I_n^{-1}(Y-XB)'\Sigma^{-1}(Y-XB)}$$

It can be shown that

$$\hat{B} = (X'X)^{-1}X'Y$$

and

$$\hat{\Sigma} = \frac{1}{n}(Y - X\hat{B})'(Y - X\hat{B})$$

are MLE's.

If $Y = (Y_1, Y_2)$ and $\hat{B} = (\hat{\alpha}, \hat{\gamma})$, then

$$\hat{\Sigma} = \frac{1}{n}\begin{bmatrix} (Y_1 - X\hat{\alpha})'(Y_1 - X\hat{\alpha}) & (Y_1 - X\hat{\alpha})'(Y_2 - X\hat{\gamma}) \\ (Y_2 - X\hat{\gamma})'(Y_1 - X\hat{\alpha}) & (Y_2 - X\hat{\gamma})'(Y_2 - X\hat{\gamma}) \end{bmatrix}$$

Two Voxel Cont.

Repeat steps to get the previous two voxel data 10000 times.

Can't show a scatter plot of the $\hat{\alpha}_0$'s, $\hat{\alpha}_1$'s, $\hat{\alpha}_2$'s, $\hat{\gamma}_0$'s, $\hat{\gamma}_1$'s, $\hat{\gamma}_2$'s.

Hexavariate distribution.

Two Voxels Cont.

It can be shown that

$$\underset{3\times 2}{\hat{B}} \sim T(n-2-1, B, (X'X)^{-1}, \tilde{\Sigma})$$

$$p(\hat{B}|\cdot) = \frac{k_T |\tilde{\Sigma}|^{\frac{n-2-1}{2}} |(X'X)^{-1}|^{-\frac{2}{2}}}{\left|\tilde{\Sigma} + \frac{1}{n-2-1}(\hat{B}-B)'[(X'X)^{-1}]^{-1}(\hat{B}-B)\right|^{\frac{(n-2-1)+3}{2}}}$$

where

$$\tilde{\Sigma} = \frac{1}{n-2-1}(Y - X\hat{B})'(Y - X\hat{B}).$$

Two Voxels Cont.

Compare to the previous multivariate t-distribution

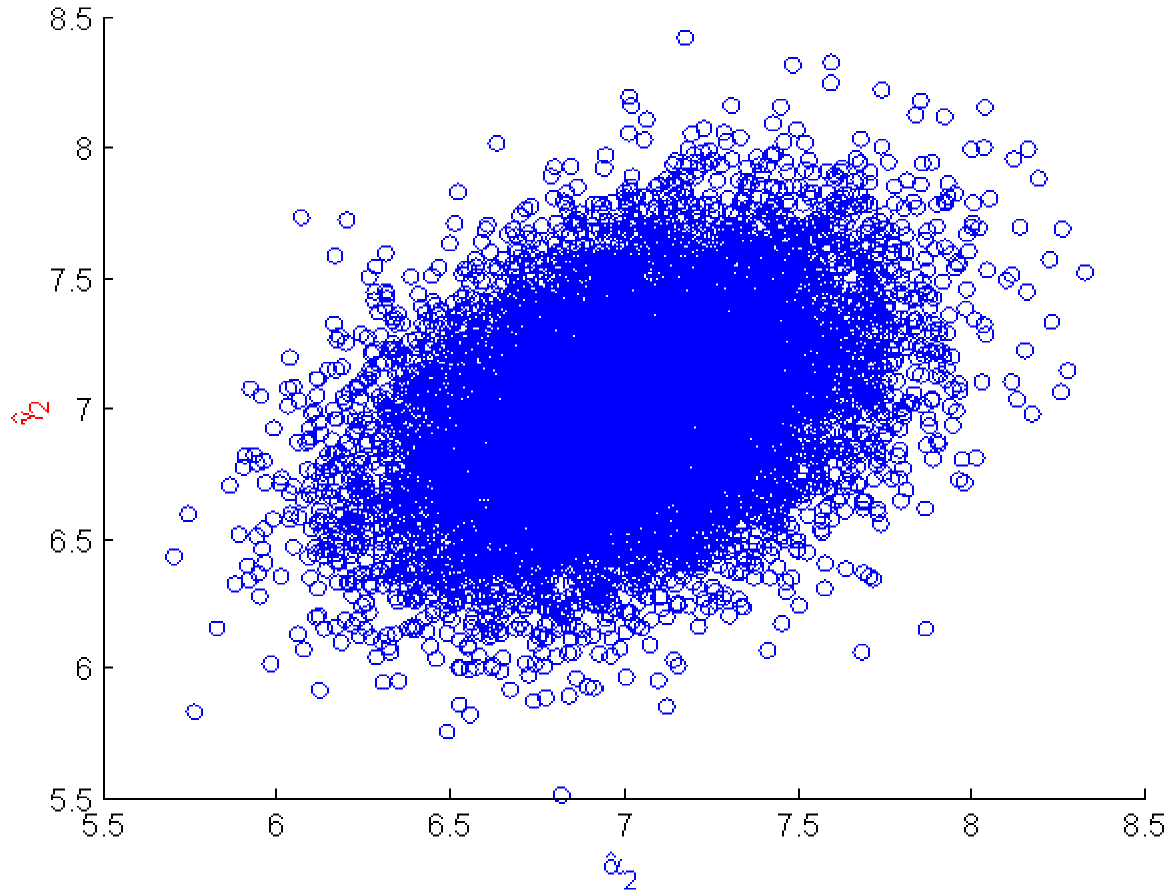
$$\underset{p \times n}{T} \sim T(\nu, T_0, \Sigma, \Phi)$$

$$p(T|\nu, T_0, \Sigma, \Phi) = k_T \frac{|\Phi|^{\frac{\nu}{2}}|\Sigma|^{-\frac{n}{2}}}{|\Phi + \frac{1}{\nu}(T - T_0)'\Sigma^{-1}(T - T_0)|^{\frac{\nu+p}{2}}}$$

Two Voxel Cont.

Scatter plot of the $\hat{\alpha}_2$ and $\hat{\gamma}_2$'s.

Gravity in the $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\gamma}_0$, and $\hat{\alpha}_1$ directions.



Bivariate distribution. Histogram.

Two Voxels Cont.

We can compute the marginal distribution of $\hat{\beta}_2 = (\hat{\alpha}_2, \hat{\gamma}_2)'$

$$p(\hat{\alpha}_2, \hat{\gamma}_2) = \int\int\int\int p(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2)\, d\hat{\alpha}_0\, d\hat{\alpha}_1 d\hat{\gamma}_0\, d\hat{\gamma}_1.$$

It can be shown that

$$\underset{2\times 1}{(\hat{\alpha}_2, \hat{\gamma}_2)'} \sim t(n-2-2, (\alpha_2, \gamma_2)', \tilde{\Sigma}, 1)$$

$$p(\hat{\beta}_2) = \frac{k_t \left[W_{33}\tilde{\Sigma}\right]^{-\frac{1}{2}}}{\left\{1 + \frac{1}{n-2-2}\left(\hat{\beta}_2 - \beta_2\right)' \left[W_{33}\tilde{\Sigma}\right]^{-1}\left(\hat{\beta}_2 - \beta_2\right)\right\}^{\frac{(n-2-2)+2}{2}}}$$

where $W = (X'X)^{-1}$ and W_{33} is the 33 element.

$$E(\hat{\beta}_2) = \beta_2$$

$$var(\hat{\beta}_2) = \frac{(n-2-2)}{(n-2-2)-2} W_{33}\tilde{\Sigma}$$

Two Voxels Cont.

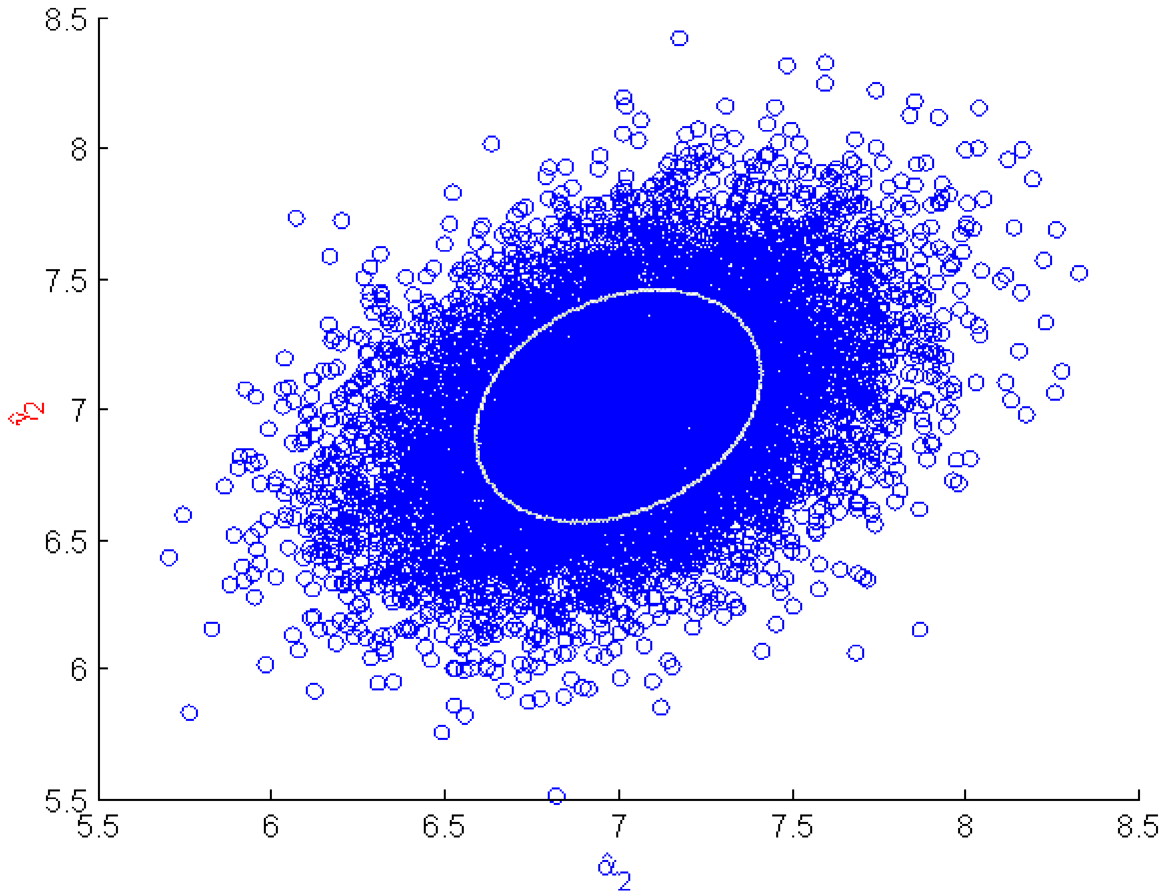
Compare to the previous student t-distribution

$$\underset{p \times 1}{t} \sim t \sim t(\nu, t_0, \Sigma, \phi^2)$$

$$p(t|\nu, t_0, \Sigma, \phi^2) = k_t \frac{(\phi^2)^{\frac{\nu}{2}} |\Sigma|^{-\frac{1}{2}}}{[\phi^2 + \frac{1}{\nu}(t-t_0)'\Sigma^{-1}(t-t_0)]^{\frac{\nu+p}{2}}}$$

Two Voxel Cont.

Scatter plot of the $\hat{\alpha}_2$ and $\hat{\gamma}_2$'s.

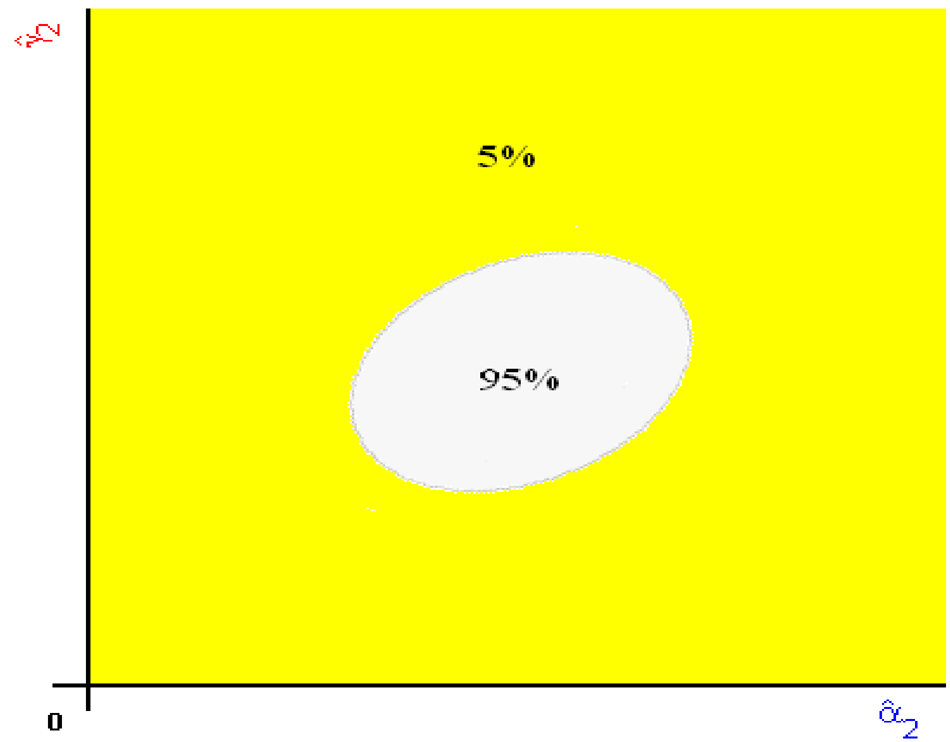


Bivariate distribution. Histogram.

Lower a parallel plane to get an ellipse.

A $(1-\alpha)*100\%$ CI for $\hat{\alpha}_2$ & $\hat{\gamma}_2$ in the ellipse.

Two Voxel Thresholding



Mound or hill which is the bivariate distribution.

Lower a plane with 95% inside ellipse or 5% outside.

Declare voxel pairs outside the ellipse active.

Hypothesis test of $(\hat{\alpha}_2, \hat{\gamma}_2) = (\alpha_2, \gamma)'$.

Two Voxel Thresholding

Geometry Review:

We learned the equation of an ellipse

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1.$$

The general form of an ellipse (tilted or not tilted) is

$$ax^2 + by^2 + cxy + dx + ey + f = 0, \qquad c^2 - 4ab < 0.$$

If $c = 0$ then not tilted.

Two Voxel Thresholding

An F-statistic (Rowe Brownbag, paper in submission.)

$$F = \frac{(n-2-2)}{2} W_{33}^{-1} (\hat{\beta}_2 - \beta_2)' G^{-1} (\hat{\beta}_2 - \beta_2)$$

$$= \frac{1}{2W_{33}} (\hat{\beta}_2 - \beta_2)' \Lambda (\hat{\beta}_2 - \beta_2) \qquad \Lambda = [G/(n-2-2)]^{-1}$$

$$= \frac{1}{2W_{33}} (\hat{\alpha}_2, \hat{\gamma}_2) \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} \begin{pmatrix} \hat{\alpha}_2 \\ \hat{\gamma}_2 \end{pmatrix} \qquad \Lambda_{12} = \Lambda_{21}$$

$$= \frac{1}{2W_{33}} (\Lambda_{11}\hat{\alpha}_2^2 + 2\Lambda_{12}\hat{\alpha}_2\hat{\gamma}_2 + \Lambda_{22}\hat{\gamma}_2^2) \qquad \leftarrow \text{an ellipse}$$

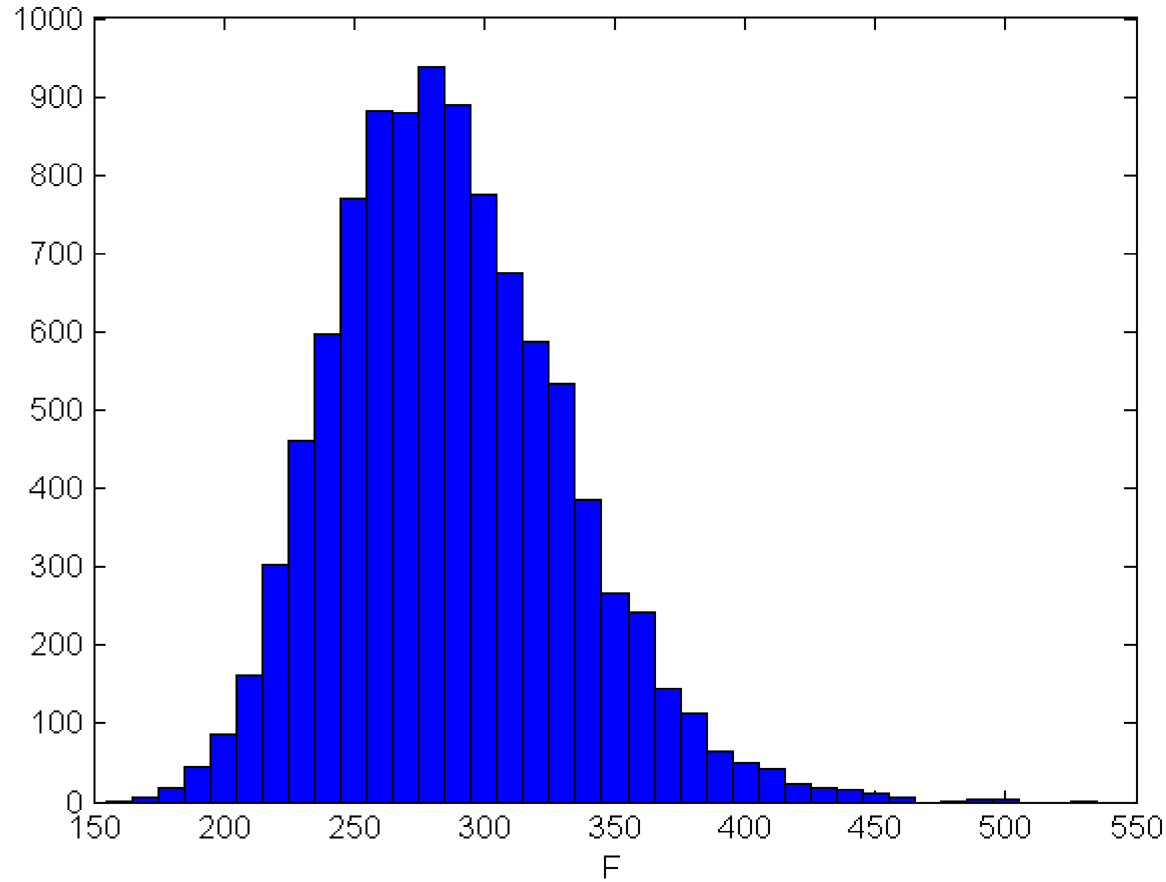
A significance level which determines F and an ellipse.

$$= \frac{(124)}{2(0.0079)} (7.3038, 6.6355) \begin{pmatrix} .00198 & .00048 \\ .00048 & .00088 \end{pmatrix} \begin{pmatrix} 7.3038 \\ 6.6355 \end{pmatrix} = 316.88$$

Look at F-distribution $F(2, n-2-2)$ and if the test statistic is in the upper or lower tail with a certain probability reject.

$F = t't$, where $t = (pW_{33}\tilde{\Sigma})^{-1/2}(\hat{\beta}_2 - \beta_2)$

Two Voxel Thresholding



Histogram of the F's.

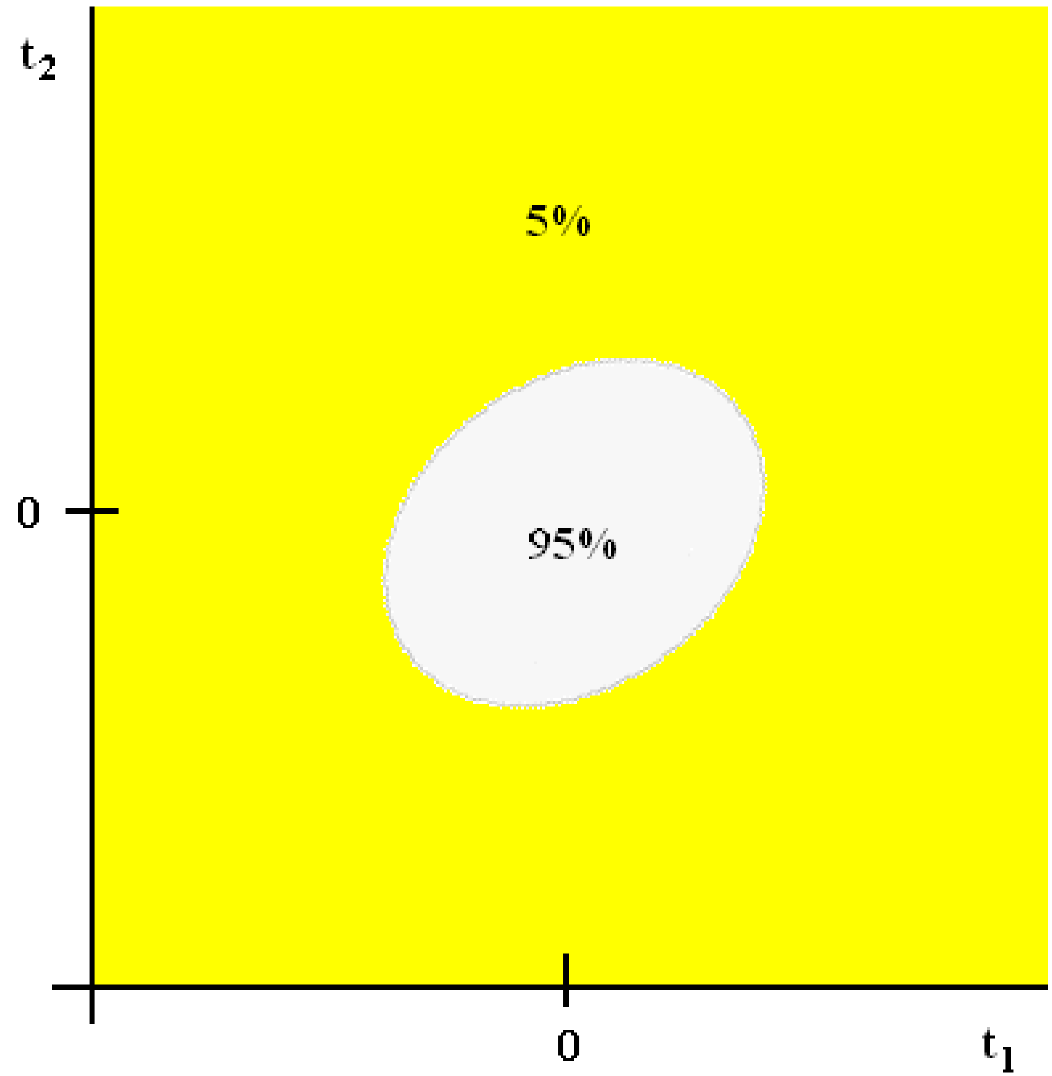
	mean	variance
Hist	287.2	1997.1
True	287.5	1960

Two Voxel Thresholding

Actually consider the joint t-distribution of the two t statistics.

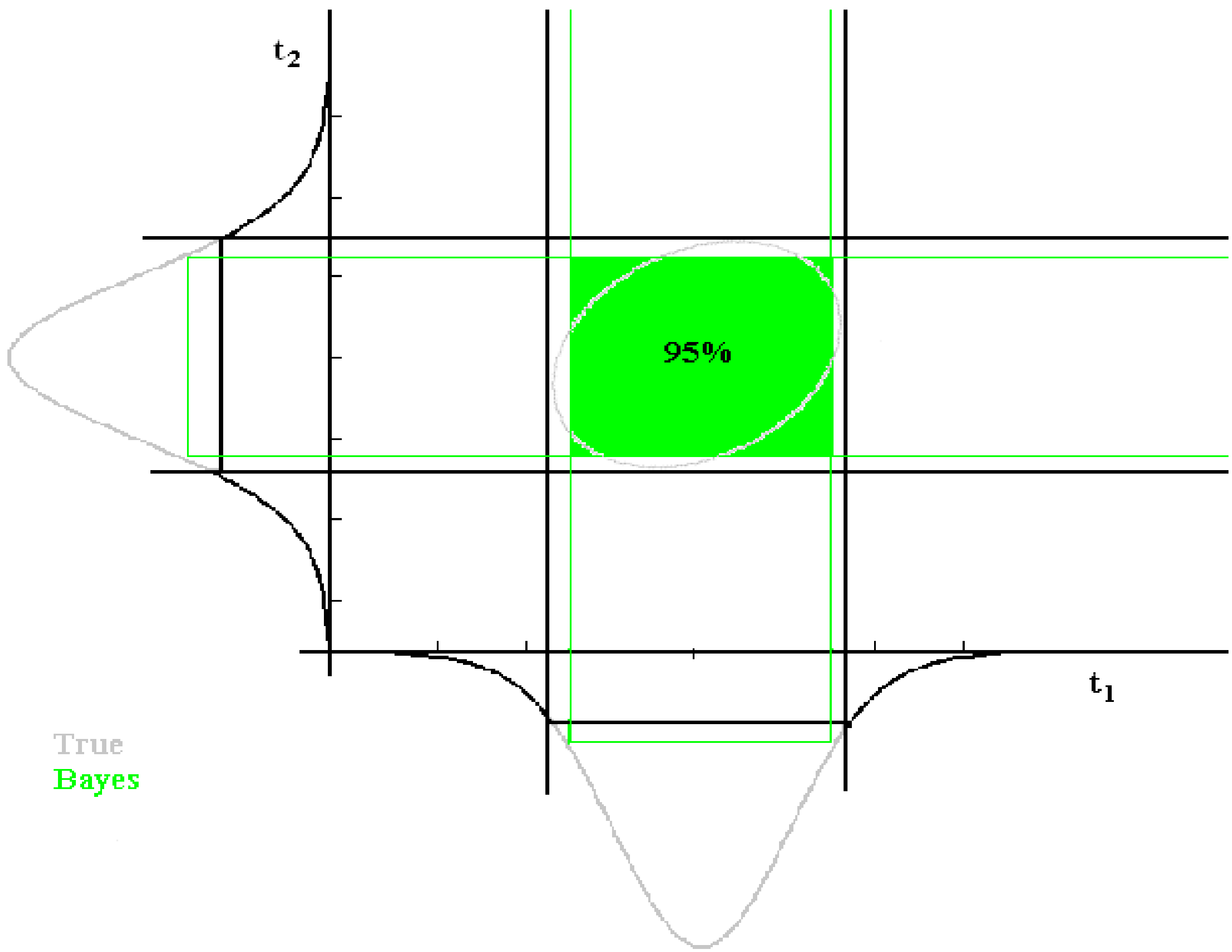
$$t_1 = \frac{\hat{\alpha}_2 - 0}{\sqrt{W_{33}\tilde{\sigma}_1^2}}$$

$$t_2 = \frac{\hat{\gamma}_2 - 0}{\sqrt{W_{33}\tilde{\sigma}_2^2}}$$



Two Voxel Thresholding

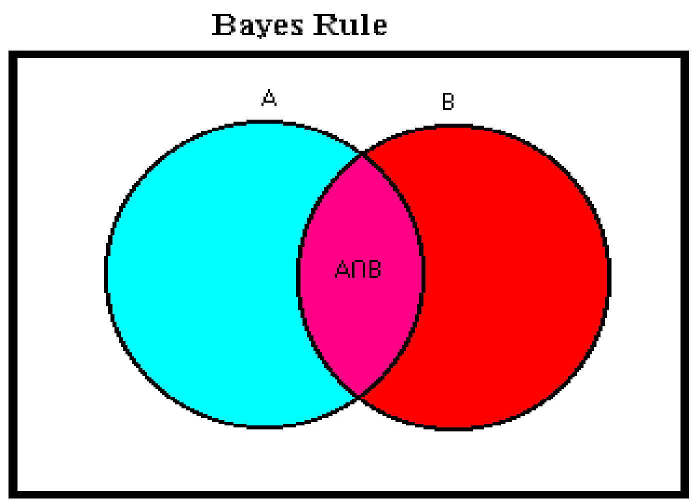
Without the ellipse consider a rectangle.
We determine thresholds based on marginal tail areas.



Two Voxel Thresholding

Thresholds based on marginal tail areas.

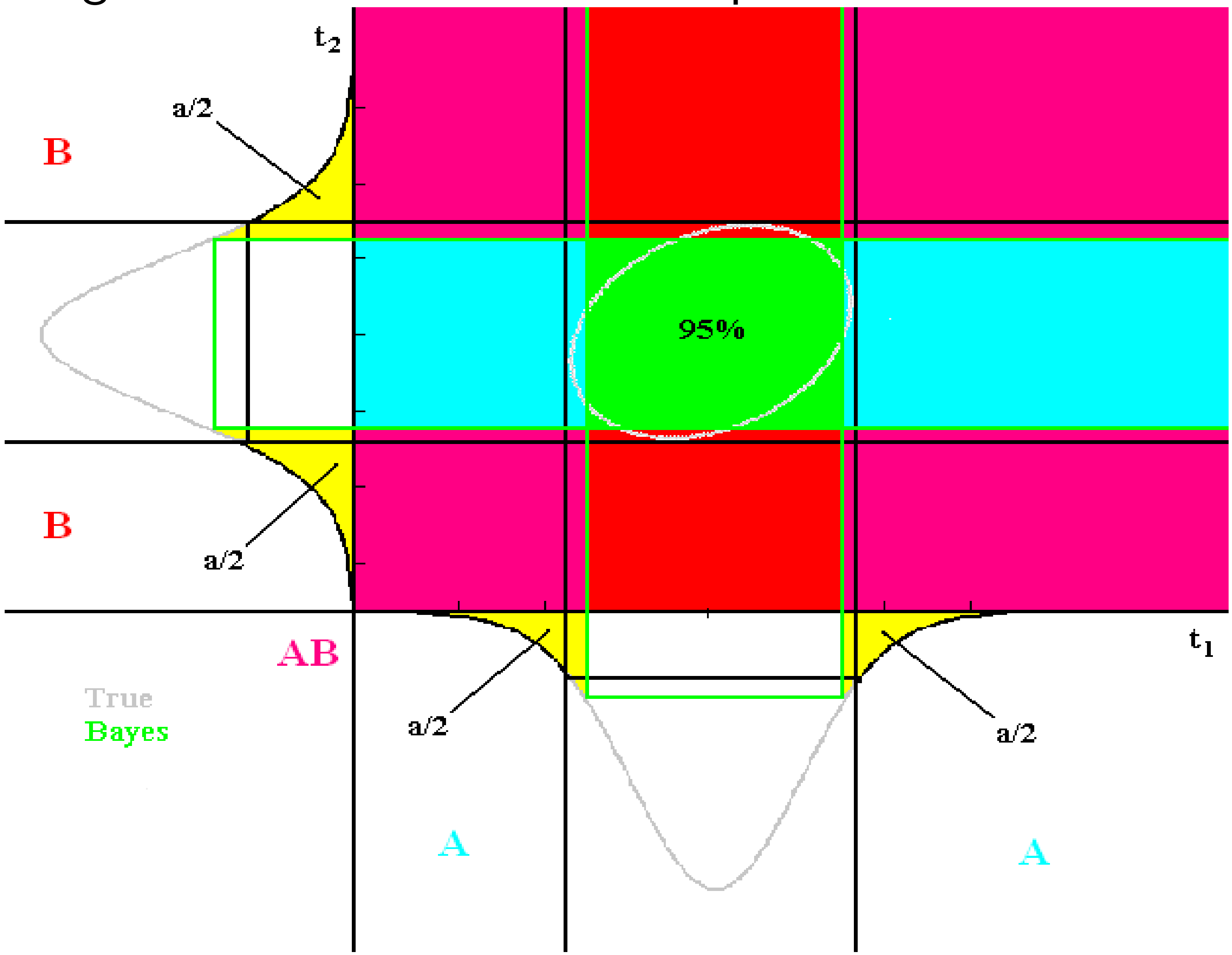
$$
\begin{aligned}
BAY &= P(|t_1| > c_1) + P(|t_2| > c_2) - P(|t_1| > c_1 \cap |t_2| > c_2) \\
&= a + a - a' \\
&= \alpha
\end{aligned}
$$



Assuming we know $P(|t_1| > c_1 \cap |t_2| > c_2)$.
i.e. The joint distribution.

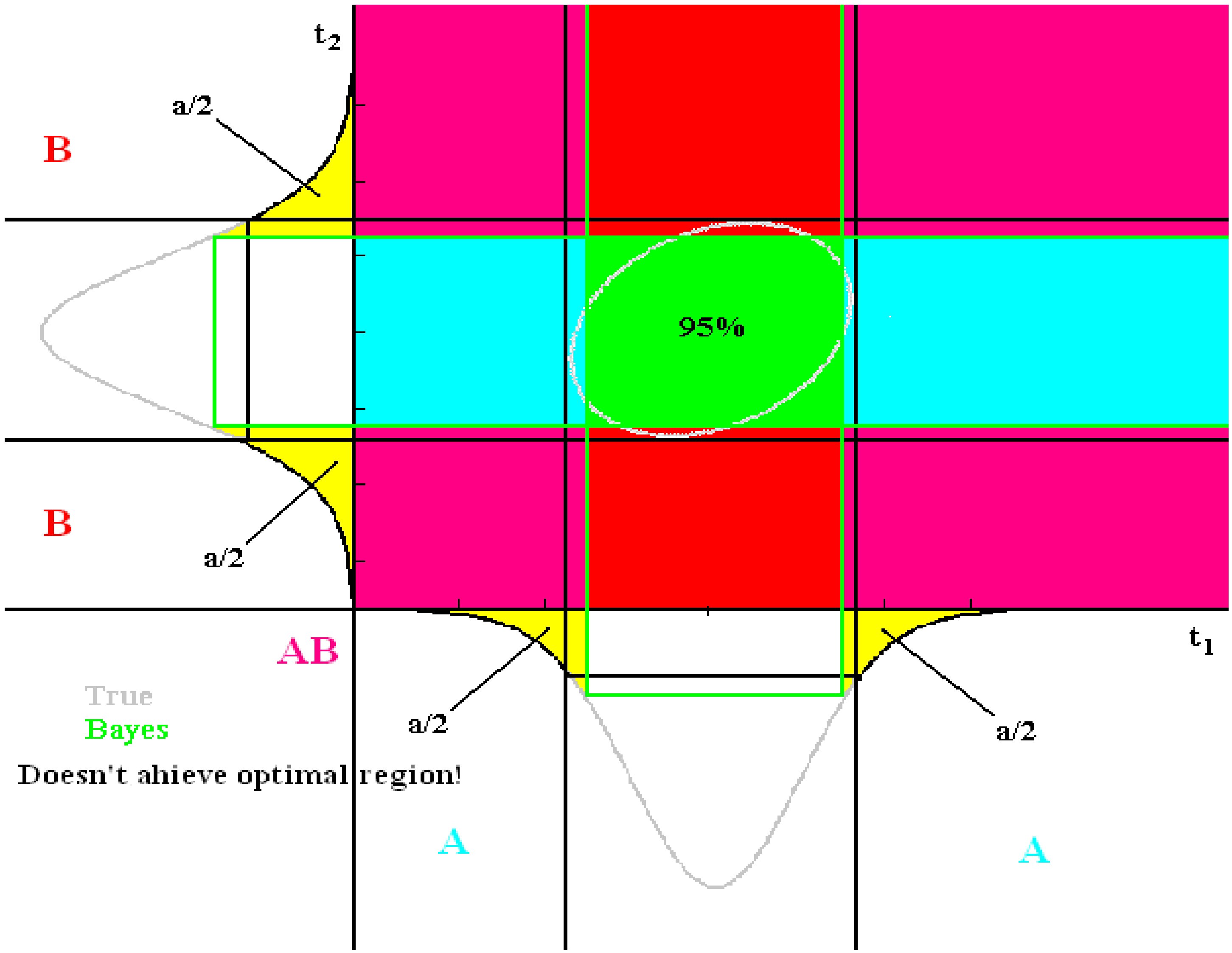
Two Voxel Thresholding

Same as putting $a/2$ in the tails of the marginals.
The rectangle is not the same as the ellipse.



Two Voxel Thresholding

We put some significance value in each of the yellow tails.
Less area in the tails means more area in the rectangle and vise versa.



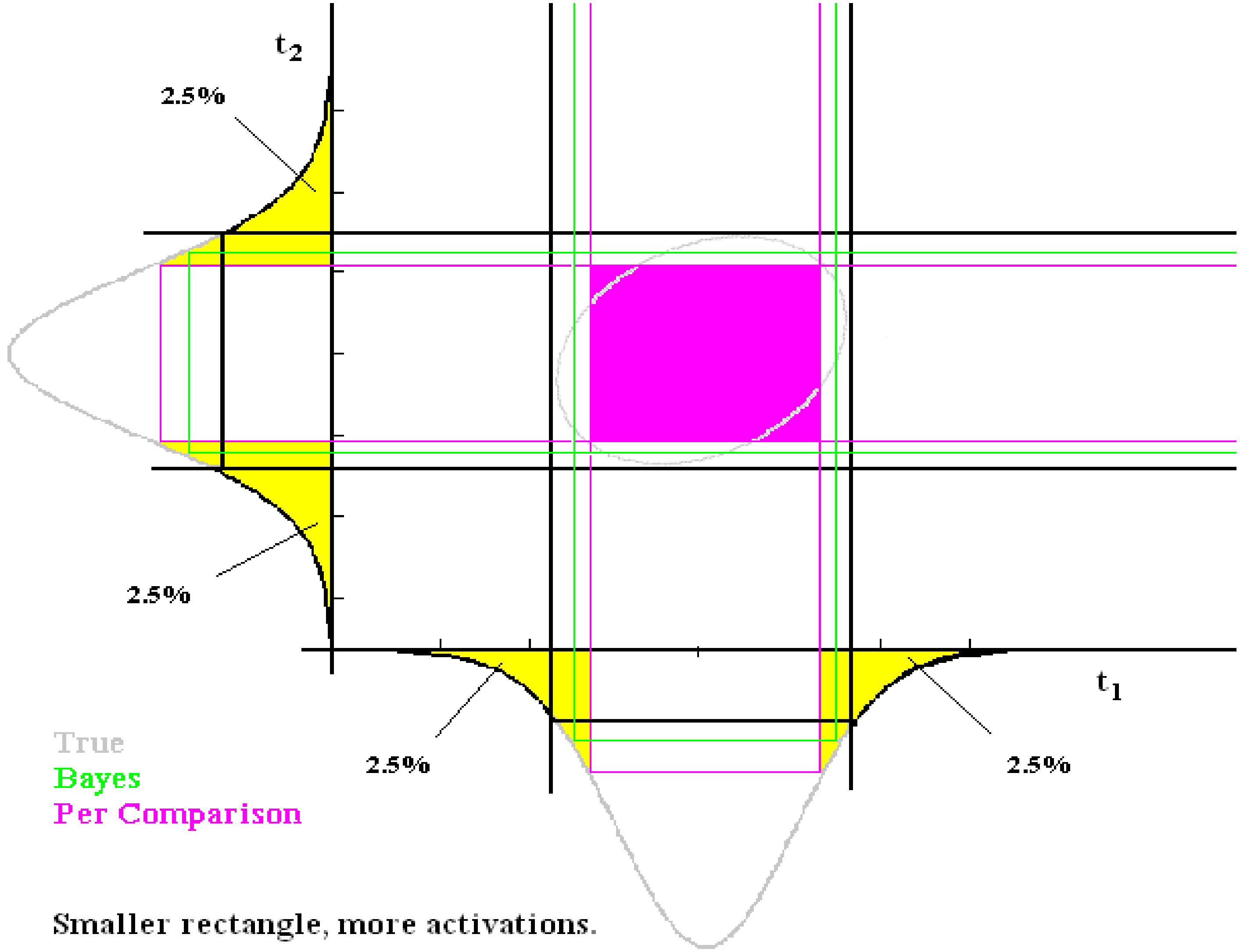
Two Voxel Thresholding

Assuming we don't know $P(|t_1| > c_1 \cap |t_2| > c_2)$.
No one ever uses the joint distribution!
Everyone uses the marginals!

$$
\begin{aligned}
BAY &= P(|t_1| > c_1) + P(|t_2| > c_2) - P(|t_1| > c_1 \cap |t_2| > c_2) \\
&= 2a \qquad \left(\frac{a}{2}\,\text{in each tail}\right) \\
PCE &= P(|t_1| > c_1) + P(|t_2| > c_2) \\
&= 2\alpha \qquad \left(\frac{\alpha}{2}\,\text{in each tail}\right) \\
BON &= P(|t_1| > c_1) + P(|t_2| > c_2) \\
&= \alpha \qquad \left(\frac{\alpha}{2p}\,\text{in each tail}\right)
\end{aligned}
$$

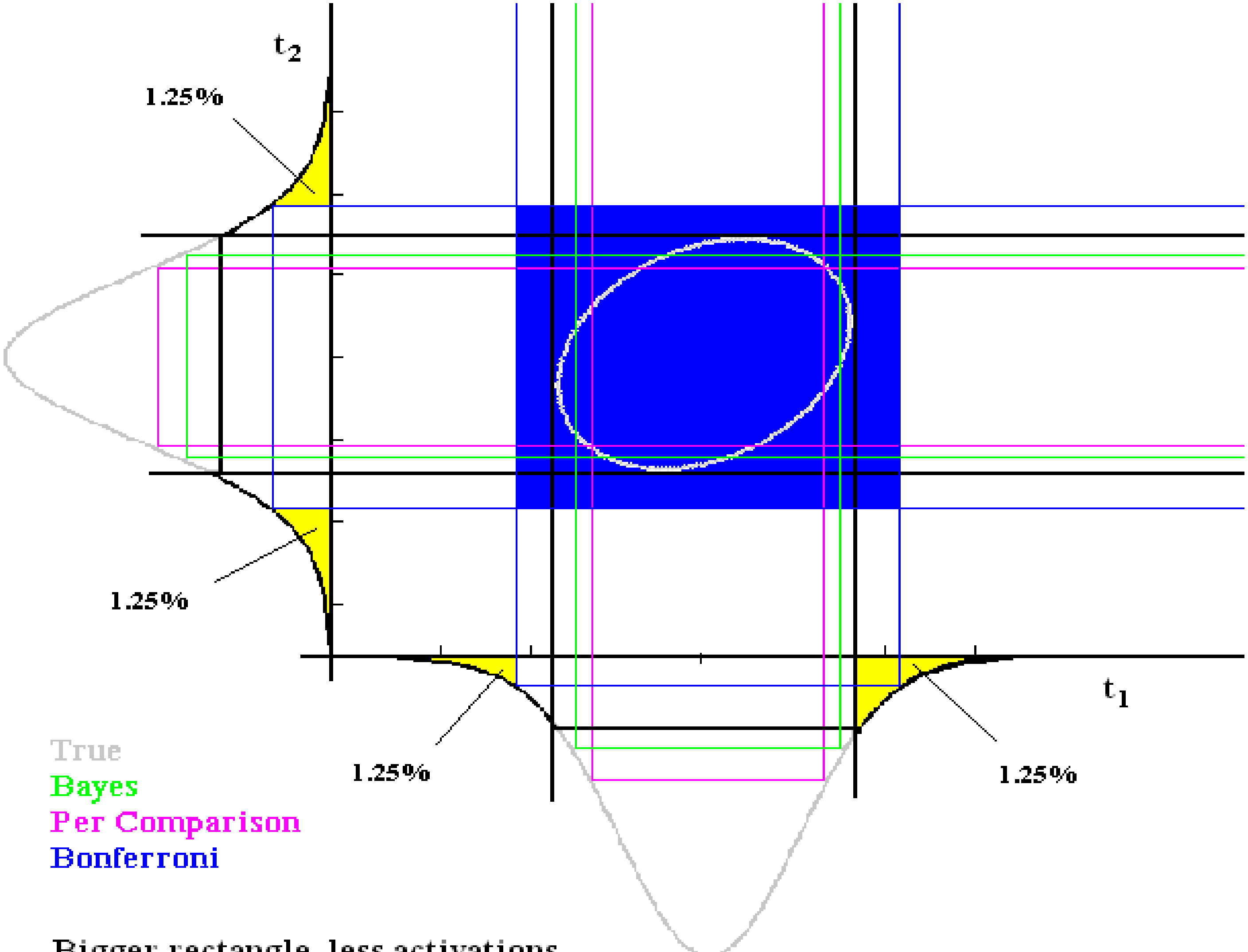
$$\alpha < 2a < 2\alpha$$

Two Voxel Thresholding



Smaller rectangle, more activations.

Two Voxel Thresholding



Bigger rectangle, less activations.

Two Voxel Thresholding

What about a hypothesis test to see if Σ is diagonal?
(Rowe, 2003. MCMA Vol. 9 No. 4.)

That is, are the two voxels independent? Use the test statistic

$$v = -[\nu - (2p + 5)/6] \ln |\hat{R}|.$$

Monte Carlo critical values by Rowe or asymptotic χ^2 distribution with $p(p-1)/2$ DF where $\nu = n - q - p$.

Remarks

Multivariate statistics can be useful!

Catch:
The F statistic ellipse is only computable when $p \leq n$.
When $p \leq n$, G is positive definite and G^{-1} is computable.
The F (joint) statistic says, yes they are together active.
The marginal t's say yes, this one is active, and this one, etc.

Solution:
Pseudoinverse with SVD? G is at least nonnegative definite.
But the derivation assumed that G^{-1} was computable.
Bayesian approach.

Other Options: FDR, and other resampling methods.

Future Research: Singular Normal & T distributions.