Cochran's Theorem, Proof taken in large part from Gut pg 139.

- Quadratic forms are important
  - least squares
  - ANOVA
  - regression
- Basic idea: split sum of squares into a number of quadratic forms, each of which corresponds to some cause of variation
  - example: crop yield vs
    - a) fertilizer concentration
    - b) amount of water and irrigation
    - c) units of sunlight
    - d) etc
- Result: each quadratic form corresponds to one cause with one final form, the residual form, that measures the random errors involved in the experiment

- Cochran's thm. says that all of these quadratic forms are independent and χ^2 distributed

- This can be used to test hypotheses about the effects of (or influence of) different inputs on the output

The multivariate Gaussian exponent:

Thm: Suppose that $\vec{x} \in N(\vec{\mu}, \Lambda)$ with $\det(\Lambda) > 0$. Then

$$(\vec{x}-\vec{\mu})' \Lambda^{-1} (\vec{x}-\vec{\mu}) \in \chi^2(n),$$

where n is the length of $\vec{x}$.

Proof: Set $\vec{y} = \Lambda^{-1/2}(\vec{x}-\vec{\mu})$ then

$$\mathbb{E}(\vec{y}) = \mathbb{E}\left(\Lambda^{-1/2}(\vec{x}-\vec{\mu})\right) = \Lambda^{-1/2}\left(\mathbb{E}(\vec{x}) - \mu\right) = \vec{0}$$

and

$$\begin{aligned} \operatorname{cov}(\vec{y}) &= \operatorname{cov}\left(\Lambda^{-1/2}(\vec{x}-\vec{\mu})\right) \\ &= \operatorname{cov}\left(\Lambda^{-1/2}\vec{x}\right) \\ &= \Lambda^{-1/2} \operatorname{cov}(\vec{x}) \Lambda^{-1/2} \\ &= \Lambda^{-1/2} \Lambda \Lambda^{-1/2} \\ &= I \end{aligned}$$

that is, $\vec{y} \sim N(\vec{0}, I)$ and

$$\begin{aligned} (\vec{x}-\vec{\mu})' \Lambda^{-1} (\vec{x}-\vec{\mu}) &= \left(\Lambda^{-1/2}(\vec{x}-\vec{\mu})\right)' \left(\Lambda^{-1/2}(\vec{x}-\vec{\mu})\right) \\ &= \vec{y}'\vec{y} \sim \chi^2(n) \end{aligned}$$

Note, this means that the Y_i's are <u>independent</u>!

Remember

What is $A^{-1/2}$?

Assume A is a symmetric matrix (all quadratic forms are symmetric) and recall that symmetric matrices can be diagonalized

$$C'AC = D$$

where D is a diagonal matrix and C is an orthogonal matrix, i.e., $C^{-1} = C' \Leftrightarrow C'C = I$.

The diagonal elements of D are the eigenvalues, $\lambda_1 \ldots \lambda_n$ of A.

Clearly $\det A = \det D = \prod_{k=1}^{n} \lambda_k$ and

$\text{trace } A = \text{trace } (CDC') = \text{trace } (D)$

$\tilde{D} = D^{1/2}$ is easy to define, it's just the $\hat{D} = [+\sqrt{d_{ij}}]$. If we set

$$B = C\hat{D}C'$$

then we have

$$B^2 = BB = C\tilde{D}C'C\tilde{D}C' = C\tilde{D}\tilde{D}C' = CDC' = A$$

so B is the square root of A, i.e. $B = A^{1/2}$.

Additionally, $(A^{-1})^{1/2} = (A^{1/2})^{-1}$

$A^{-1} = (C'DC)^{-1}$

$C'D^{-1}C(C'DC) = I$

so $(C'DC)^{-1} = C'D^{-1}C$

same sqrt arg holds

To introduce Cochran's thm: consider

$\vec{X} \sim N(\vec{0}, \sigma^2 I)$, denote $\overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$

consider

$$\sum_{k=1}^{n} X_k^2 = \sum_{k=1}^{n} (X_k - \bar{X})^2 + n \overline{X}_n^2$$

we know that

$$\sum_{k=1}^{n} (x_k - \bar{x})^2 = (n-1) s_n^2$$ where s_n^2 is the sample variance. We know ~~from~~ prev. stats classes that

$$(n-1) s_n^2 \sim \sigma^2 \chi^2(n-1)$$

We also know that

$$n \overline{X}_n^2 \sim \sigma^2 \chi^2(1)$$

and

$$\sum_{k=1}^{n} (X_k)^2 \sim \sigma^2 \chi^2(n)$$

$x_k \sim N(0, \sigma^2)$

$\vec{X}^T I X \sim ? \chi^2(n)$

$\frac{X_k}{\sigma} \sim N(0, 1)$

$(\frac{1}{\sigma} X)^T (\frac{1}{\sigma} X) \sim \chi^2(n)$

$X^T X \sim \sigma^2 \chi^2(n)$

If we let $\vec{y} = \frac{1}{\sigma}\vec{X}$ then $\vec{y} \sim N(0, I)$ and we can prove $\vec{y}^T y \sim \chi^2(n)$ by method of moment generating functions.

We know $y_1^2 \sim \chi^2(1)$ $\Bigg\{$

What is the distribution of

$\sum_{k=1}^{n} y_k^2 \sim ?$

We know that the distribution of a sum of random variables is the product of their moment generating functions.

Definition Let X be a random variable. The MGF of X is

$$\psi_X(t) = E(e^{tX})$$

Thm Let $X_1, X_2, \ldots X_n$ be independent r.v.'s whose MGF's exist. Set $S_n = X_1 + X_n + \cdots X_n$ then

$$\psi_{S_n}(t) = \prod_{k=1}^{n} \psi_{X_k}(t)$$

M.G.F. for $\chi^2(v)$ is $\psi_{\chi^2}(t) = (1-2t)^{v/2}$ (Wackerly)

$$\psi_{S_n}(t) = \prod_{k=1}^{n} (1-2t)^{v/2} = (1-2t)^{nv/2} \Rightarrow S_n \sim \chi^2(n)$$

But.

$$\sum_{k=1}^{n} X_k^2 = \sum_{k=1}^{n} (x_k - \bar{X}_n)^2 + n \bar{X}_n^2$$

is equivalent to

$$\vec{x}' I \vec{x} = \vec{x}'\left(I - \tfrac{1}{n} J\right)\vec{x} + \tfrac{1}{n} \vec{x}' J \vec{x}.$$

and from this we can arrive at the fact that

$\frac{1}{\sigma^2} \vec{x}' I \vec{x} \sim \chi^2(n)$ if $x_i \sim N(0, \sigma^2)$ using Cochran's Thm easily.

Toward's proving Cochran's Thm. we start with the following lemma (Gut)

Let $x_1, x_2, \cdots x_n$ be real numbers. Suppose that $\sum_{i=1}^{n} x_i^2$ can be split into a sum of non-negative definite quadratic forms, that is,

$$\sum_{i=1}^{n} x_i^2 = Q_1 + Q_2 + \cdots + Q_k,$$

where $Q_i = x' A_i x$ and $\operatorname{rank} Q_i \equiv \operatorname{rank}(A_i) = r_i \; \forall i$.

If $\sum_{i=1}^{k} r_i = n$ then $\exists$ an <u>orthogonal</u> matrix C s.t.

$$x = Cy$$

a-d

$$Q_1 = Y_1^2 + Y_2^2 + \cdots + Y_{r_1}^2$$

$$Q_2 = Y_{r_1+1}^2 + Y_{r_1+2}^2 + \cdots + Y_{r_1+r_2}^2$$

$$Q_3 = Y_{r_1+r_2+1}^2 + \cdots \quad Y_{r_1+r_2+r_3}^2$$

$$Q_k = Y_{n-r_k+1}^2 + Y_{n-r_k+2}^2 + \cdots \quad Y_n^2$$

Note: each of the quad forms contains different, non-overlapping sets of y_i's, and that the # of terms in each Q_i is r_i

We start with the case $n=2$. The general case can be proved by induction.

Proof for $n=2$: By assumption we have

$$Q = \sum_{i=1}^{n} x_i^2 = x' A_1 x + x' A_2 x = Q_1 + Q_2$$

where A_1 and A_2 are non-negative definite (symmetric: they are quadratic forms) matrices with ranks r_1 and r_2. ($r_1 + r_2 = n$ also by assumption).

note we know that A is symmetric, if A is positive semi-definite iff $\vec{b}' A \vec{b} \geq 0 \quad \forall b \neq 0$

We know that symmetric matrices can be orthogonally diagonalized, ie $C'AC = \Lambda$. If A is pos. semi. def. then

$$b' \Lambda b = b'(C'AC)b = \tilde{b}' A \tilde{b} \geq 0$$

where $\tilde{b} = Cb$. This means that Λ is also pos. semi. definite which means that Λ must have all positive entries (A has all positive eig. v's)

Because A_1 is symmetric it can be orthogally diagonalized (check, Lin Alg. review pg 33)

$$C'A_1C = D$$

where D is diagonal, and the diagonal elements of which are the eigen values of A_1. Since rank $(A_1) = r_1$ we know that rank$(A_1) = \#$ non-zero eigen values in D (non-singular transforms preserve rank, ie. rank(BA) = rank(A) if B non singular, rank is number of linearly independent col's or row's)

Since rank$(A_1) = r_1$, and A_1 is symmetric then we have $\lambda_1, \ldots, \lambda_{r_1}$ positive eigen values and $n - r_1$ λ-values equal to zero.

Set $x = Cy$

$$Q = \sum_{i=1}^{n} x_i^2 = x'Ix = y'C'ICy = y'y = \sum y_i^2$$

$$= y'C'A_1Cy + y'C'A_2Cy$$

$$= y'Dy + y'C'A_2Cy$$

$$= \sum_{i=1}^{r_1} \lambda_i y_i^2 + y'C'A_2Cy$$

Equivalently we can write

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{r_1} \lambda_i y_i^2 + y' C' A_2 C y$$

$$\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{r_1} \lambda_i y_i^2 = y' C' A_2 C y$$

$$\sum_{i=1}^{r_1} (1-\lambda_i) y_i^2 + \sum_{i=r_1+1}^{n} y_i^2 = y' C' A_2 C y$$

Now, since the rank of $C' A_2 C = r_2$ we know that $\lambda_1 \cdots \lambda_{r_1} = 1$. Note

$$\underbrace{y' \begin{bmatrix} 1-\lambda_1 & & & & & \\ & 1-\lambda_2 & & & & \\ & & \ddots & & & \\ & & & 1-\lambda_{r_1} & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} y}_{\operatorname{rank}(?) \geq r_1,\ = r_1 \text{ iff. } \lambda_1 \cdots \lambda_{r_1} = 1} = y' \underbrace{C' A_2 C}_{\operatorname{rank}(C' A_2 C) = r_2} y$$

Now

$$Q_1 = \sum_{i=1}^{r_1} y_i^2 \quad \text{and} \quad Q_2 = \sum_{i=r_1+1}^{n} y_i^2$$

Now in the general case we can recur on the subblocks, choosing different C's for each.

Now, Cochran's Thm is almost immediate.

C.T.: Let $X_1, X_2, \dots, X_n$ be independent $N(0, \sigma^2)$ r.v.'s and suppose

$$\underset{\sim}{x}' I \underset{\sim}{x} = \sum_{i=1}^{n} X_i^2 = Q_1 + Q_2 + \dots + Q_k$$

where $Q_1, Q_2, \dots, Q_k$ are non-negative def. quadratic forms in the X's. That is:

$$Q_i = \underset{\sim}{x}' A_i x \quad , \quad i = 1, \dots, k$$

Set rank $A_i = r_i$

$$\text{If} \quad r_1 + r_2 + \dots \, r_k = n$$

then:

a) $Q_1, Q_2, \dots Q_k$ are independent, and

b) $Q_i \sim \sigma^2 \chi^2(r_i)$, $i = 1, \dots, k$

By the lemma $\exists$ an orthogonal matrix C s.t.

$X = CY$ yields

$$Q_1 = Y_1^2 + \dots \, Y_{r_1}^2$$

$$Q_2 = Y_{r_1+1}^2 + \dots \, Y_{r_1+r_2}^2$$

$$Q_k = Y_{n-r_k+1}^2 + \dots \, Y_n^2$$

Where $Y_1, \dots Y_n$ are independent $N(0, \sigma^2)$ r.v.'s and each occurs only in one term.

Check $\vec{x} \sim N(0, \sigma^2 I)$

$$\vec{x} = C\vec{y} \qquad E(\vec{y}) = E(C'\vec{x}) = \vec{0}$$

$$Cov(\vec{y}) = Cov(C'\vec{x}) = C\, cov(X)\, C'$$

$$= \sigma^2 C I C' = \sigma^2 I$$

Now for regression. We have

$$SSTO = Y'\left[I - \left(\tfrac{1}{n}\right)J\right]Y$$

$$SSE = Y'\left[I - H\right]Y$$

$$SSR = Y'\left[H - \left(\tfrac{1}{n}\right)J\right]Y$$

and $Y'IY = Y'\left[I - \left(\tfrac{1}{n}\right)J\right]Y + Y'\left(\tfrac{1}{n}J\right)Y$

$$SSTO = SSE + SSR$$

$$Y - (X\beta) \sim N(0, \sigma^2 I)$$

$$(Y - X\beta)' I (Y - X\beta)$$

$$= Y'IY - Y'IX\beta - (X\beta)'IY + (X\beta)'(X\beta)$$