Cochran's Theorem. Proof, take 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, US
    a) fertility 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 theorem says that all of these quadratic forms are independent and $\chi^2$ distributed.

  - This can be used to test hypotheses about the effects of (or influence of) different inputs on the output.
The multivariate Gaussian exponent:

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

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

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

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

$$\mathbf{E}(\mathbf{y}) = \mathbf{E}(\Lambda^{-1/2} (\mathbf{x} - \mathbf{\mu})) = \Lambda^{-1/2} (\mathbf{E}(\mathbf{x}) - \mathbf{\mu}) = \mathbf{0},$$

and

$$\text{cov} (\mathbf{y}) = \text{cov} (\Lambda^{-1/2} (\mathbf{x} - \mathbf{\mu})).$$

$$= \Lambda^{-1/2} \text{cov} (\mathbf{x}) \Lambda^{-1/2}$$

$$= \Lambda^{-1/2} \mathbf{\Lambda} \Lambda^{-1/2}$$

$$= \mathbf{I},$$

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

$$(\mathbf{x} - \mathbf{\mu})' \mathbf{\Lambda}^{-1} (\mathbf{x} - \mathbf{\mu})$$

$$= (\Lambda^{-1/2} (\mathbf{x} - \mathbf{\mu}))' (\Lambda^{-1/2} (\mathbf{x} - \mathbf{\mu}))$$

$$= (\mathbf{y})' (\mathbf{y}) \sim \chi^2 (n).$$

Note: this means that the $y_i$'s are independent.
Remember

What is $A^{-\frac{1}{2}}$?

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., $C^{-1}CC' = I$.

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

Clearly $\det A = \det D = \prod_{i=1}^{n} \lambda_i$ and

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

$D = D^\frac{1}{2}$ is easy to define; it's just the diagonal. If we set

$$B = C D C'$$

then we have

$$B^2 = BB = C D C' C D C' = C D D C' = C D C' = A$$

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

Additionally

$$A^\frac{1}{2} = (C' D C)^{\frac{1}{2}}$$

$$\left(A^{-\frac{1}{2}}\right)^{\frac{1}{2}} = (A^{\frac{1}{2}})^{-\frac{1}{2}}$$

so $(C' D C)^{-1} = C' D^{-1} C$ some sort argholds
To introduce Cochran's test, consider

\[ \hat{X} \sim N(\theta, \sigma^2 \Gamma) \], denote \( \bar{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k \)

Consider

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

we know that

\[ \sum_{k=1}^{n} (X_k - \bar{X})^2 = (n-1) S_n \]

where \( S_n \) is the sample variance. We know from previous classes that

\[ (n-1) S_n^2 \sim \sigma^2 \chi^2(n-1) \]

We also know that

\[ n \bar{X}_n^2 \sim \sigma^2 \chi^2(1) \]

and

\[ \sum_{k=1}^{n} (X_k - \bar{X})^2 \sim \sigma^2 \chi^2(n) \]

\[ X_k \sim N(\theta, \sigma^2) \]

\[ X^T \Gamma X \sim \chi^2(n) \]

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

\[ \left( \frac{1}{\sigma} X \right)^T \left( \frac{1}{\sigma} X \right) \sim \chi^2(n) \]

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

We know \( y^2 \sim \chi^2(n) \)

What is the distribution of \( \sum y_k^2 \)?

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. Then

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

Then let \( X_1, X_2, \ldots, X_n \) be independent r.v.'s
whose MGF's exist. Set \( S_n = X_1 + X_2 + \ldots + X_n \) then

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

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

\[
\psi_{S_n}(t) = \prod_{k=1}^{n} (1 - 2t)^{-1/2} = (1 - 2t)^{-n/2} \Rightarrow S_n \sim \chi^2(n)\]
But
\[ \sum_{k=1}^{n} x_k^2 = \sum_{k=1}^{n} (x_k - \bar{x})^2 + \frac{1}{n} \bar{x}^2 \]
is equivalent to
\[ x'Jx = x'(I - \frac{1}{n}J)x + \frac{1}{n} x'Jx \]
and from this we can arrive at the fact that
\[ \frac{1}{n} x'Jx \sim \chi^2(n) \text{ if } x_i \sim N(0, \sigma^2) \text{ using Cochran's Theorem easily.} \]

Towards proving Cochran's Theorem, we start with the following lemma (Gur).

Let \( x_1, x_2, \ldots, x_n \) 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,
\[ \sum_{i=1}^{n} x_i^2 = Q_1 + Q_2 + \ldots + Q_k, \]

where \( Q_i = x' A_i x \) and rank \( Q_i = \text{rank} (A_i) = r_i \ A_i \).

If \( \sum_{i=1}^{k} r_i = n \) then there is an orthogonal matrix \( C \) such that
\[ x = Cy \]
\[ Q_1 = \gamma_1^2 + \gamma_2^2 + \ldots + \gamma_{r_1}^2 \]
\[ Q_2 = \gamma_{r_1+1}^2 + \gamma_{r_1+2}^2 + \ldots + \gamma_{r_1+r_2}^2 \]
\[ Q_3 = \gamma_{r_1+r_2+1}^2 + \gamma_{r_1+r_2+2}^2 + \ldots + \gamma_{r_2+r_3}^2 \]
\[ Q_4 = \gamma_{r_2+r_3+1}^2 + \gamma_{r_2+r_3+2}^2 + \ldots + \gamma_{r_3}^2 \]
Note: each of the quadrants contains different, non-overlapping sets of terms, and the number of terms in each $Q_i$ is $c_i$.

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

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

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

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 if $b^T A b \geq 0 \forall b \neq 0$.

We know that symmetric matrices can be orthogonally diagonalized, i.e., $C^T A C = \Lambda$. If $A$ is pos. semi-def. then

$$ b^T \Lambda b = \tilde{b}^T (C^T A C)b = \tilde{b}^T \tilde{A} \tilde{b} \geq 0 $$

where $\tilde{b} = Cb$. This means that $\Lambda$ is also pos. semi-definite which means that $\Lambda$ must have all positive entries ($A$ has all positive entries).
Because $A$ is symmetric it can be orthogonally diagonalized (check Lin Alg. review pg 77)

$$C' A C = D$$

where $D$ is diagonal, and the diagonal elements of which are the eigenvalues of $A$. Since rank $(A) = r$, we know that rank $(A) = \# \text{non-zero eigenvalues in } D$ (non-singular transforms preserve rank, i.e. rank $(B^T) = \text{rank } (A)$ if $B$ is non-singular, rank is number of linearly independent cols or rows)

Since rank $(A) = r$, and $A$ is symmetric then we have $2, \ldots, 2$, positive eigenvalues and $n-q$, $2$-values equal to zero.

Set $x = C y$

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

$$= y' C' A C y + y' C' A_2 C y$$

$$= y' D y + y' C' A_2 C y$$

$$= \sum_{i} y_i^2 + y' C' A_2 C y$$
Equivalently, we can write

$$
\sum_{e=1}^{n} y_e^2 = \sum_{e=1}^{n} y_{e'}^2 + \gamma' \gamma' A_{\gamma'} C_{\gamma'}
$$

$$
\sum_{e=1}^{n} y_e^2 = \sum_{e=1}^{n} y_{e'}^2 = \gamma' A_{\gamma'} C_{\gamma'}
$$

$$
\sum_{e=1}^{n} (1-\lambda_e) y_e^2 + \sum_{e=1}^{n} y_e^2 = \gamma' A_{\gamma'} C_{\gamma'}
$$

Now, since the rank of $\gamma' A_{\gamma'} C_{\gamma'} = \Gamma_e$
we know that $\lambda_1 \ldots \lambda_{\Gamma_e} = 1$. Note

$$
\begin{cases}
\gamma' & \text{if } \lambda_e = 1 \\
1-\lambda_e & \text{if } \lambda_e = 0
\end{cases}
$$

Now

$$Q = \sum_{e=1}^{n} y_e^2 \quad \text{and} \quad Q_1' = \sum_{e=1}^{n} y_{e'}^2$$

Now in the general case, we can

\[
\text{occur on the subblocks, choosing different } C'\text{'s for each.}
\]
Now, Cochran's Theorem is almost immediate.

**C.T.** Let \( X_1, X_2, \ldots, X_k \) be independent \( \mathcal{N}(0, \sigma^2) \) r.v.'s and suppose

\[
\sum_{i=1}^k X_i^2 = Q_1 + Q_2 + \ldots + Q_k,
\]

where \( Q_1, Q_2, \ldots, Q_k \) are non-negative, non-degenerate quadratic forms in the \( X_i \)'s. That is:

\[
Q_i = Y_i' A_i Y_i, \quad i = 1, \ldots, k
\]

Let \( r_1 \) be \( \begin{bmatrix} A_1 \end{bmatrix} \).

If \( r_1 + r_2 + \ldots + r_k = n \)

then:

a) \( Q_1, Q_2, \ldots, Q_k \) are i-dependent, and

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

By the lemma, let an orthogonal matrix \( C \) s.t.

\[
X = CY \quad \text{yields}
\]

\[
Q_1 = Y_1^2 + \ldots + Y_{r_1}^2
\]

\[
Q_2 = Y_{r_1+1}^2 + \ldots + Y_{r_1+r_2}^2
\]

\[
Q_k = Y_{r_1+r_2+\ldots+r_{k-1}+1}^2 + \ldots + Y_n^2
\]

Where \( Y_1, \ldots, Y_n \) are independent \( \mathcal{N}(0, \sigma^2) \) r.v.'s and each occurs only in one term.
Check \( \hat{\beta} \sim N(0, \sigma^2 I) \)

\[
\hat{\beta} = C \hat{y} \\
E(\hat{\beta}) = E(C' \hat{y}) = 0 \\
Cov(\hat{\beta}) = Cov(C' \hat{y}) = C \operatorname{Cov}(\hat{y}) C' \\
= \sigma^2 \mathbb{I} C' = \sigma^2 \mathbb{I}
\]

Now for regression. We have

\[
\begin{align*}
SS_{TOT} &= y' \left[ I - \left( \frac{1}{n} \right) J \right] y \\
SS_{E} &= y' \left[ I - H \right] y \\
SS_{R} &= y' \left[ H - \left( \frac{1}{n} \right) J \right] y \\
\end{align*}
\]

and

\[
y' I y = y' \left[ I - \left( \frac{1}{n} \right) J \right] y + \frac{1}{n} \mathbb{I} y
\]

\[
SS_{TOT} = SS_{E} + SS_{R}
\]

\[
y - (x \beta) \sim N(0, \sigma^2 \mathbb{I})
\]

\[
(y - x \beta)' \mathbb{I} (y - x \beta) \\
= y' I y - y' I x \beta - (x \beta)' I y + (x \beta)' (x \beta)
\]