Applied Statistical Methods II

Repeated measures model part I (a.k.a a gentle introduction to random effects models)

A simple repeated measures model

- Let Y_{ij} be the response for the j^{th} sample from subject i.
- Since Y_{ij}, $Y_{ij'}$ are from the same individual, they will be **correlated**. We must model this correlation.
- The single factor random effect model is:
  - $Y_{ij} = \underbrace{\mu}_{\text{Global mean}} + \underbrace{\delta_i}_{\text{subject-specific effect}} + \underbrace{\epsilon_{ij}}_{\text{random sampling error}}$
  - $\mu_i = \mu + \delta_i \sim$ i.i.d (μ, σ^2_μ)
  - $\epsilon_{ij} \sim$ i.i.d $(0, \sigma^2)$
  - μ_i's is independent of ϵ_{jk}'s
- In this simple example:
  - μ is a **fixed effect**, since it is NOT random.
  - δ_i is a **random effect**, since individuals i are assumed to be randomly drawn from some population

A simple repeated measures model

- Let Y_{ij} be the response for the j^{th} sample from subject i.
- Since Y_{ij}, $Y_{ij'}$ are from the same individual, they will be **correlated**. We must model this correlation.
- The single factor random effect model is:
  - $Y_{ij} = \underbrace{\mu}_{\text{Global mean}} + \underbrace{\delta_i}_{\text{subject-specific effect}} + \underbrace{\epsilon_{ij}}_{\text{random sampling error}}$
  - $\mu_i = \mu + \delta_i \sim$ i.i.d (μ, σ_μ^2)
  - $\epsilon_{ij} \sim$ i.i.d $(0, \sigma^2)$
  - μ_i's is independent of ϵ_{jk}'s
- In this simple example:
  - μ is a **fixed effect**, since it is NOT random.
  - δ_i is a **random effect**, since individuals i are assumed to be randomly drawn from some population

A simple repeated measures model

- Let Y_{ij} be the response for the j^{th} sample from subject i.
- Since Y_{ij}, $Y_{ij'}$ are from the same individual, they will be **correlated**. We must model this correlation.
- The single factor random effect model is:
  - $Y_{ij} = \underbrace{\mu}_{\text{Global mean}} + \underbrace{\delta_i}_{\text{subject-specific effect}} + \underbrace{\epsilon_{ij}}_{\text{random sampling error}}$
  - $\mu_i = \mu + \delta_i \sim$ i.i.d (μ, σ^2_μ)
  - $\epsilon_{ij} \sim$ i.i.d $(0, \sigma^2)$
  - μ_i's is independent of ϵ_{jk}'s
- In this simple example:
  - μ is a **fixed effect**, since it is NOT random.
  - δ_i is a **random effect**, since individuals i are assumed to be randomly drawn from some population

Random effect and fixed effect

- Fixed Effect: we are interested in the mean value for each factor level i, and relevant inferences.
- Random Effect - mean value of each level i (each subject i) is not the main interest, the levels (subjects) are viewed as a random draw from a population.
- These are just two of the many definitions of fixed and random effects...

Random effect and fixed effect

- Fixed Effect: we are interested in the mean value for each factor level i, and relevant inferences.
- Random Effect - mean value of each level i (each subject i) is not the main interest, the levels (subjects) are viewed as a random draw from a population.
- These are just two of the many definitions of fixed and random effects...

Andrew Gelman's blog

People are always asking me if I want to use a fixed or random effects model for this or that. I always reply that these terms have no agreed-upon definition. People with their own favorite definition of "fixed and random effects" don't always realize that other definitions are out there. Worse, people conflate different definitions.

Five definitions

Here are the five definitions I've seen:

(1) Fixed effects are constant across individuals, and random effects vary. For example, in a growth study, a model with random intercepts a_i and fixed slope b corresponds to parallel lines for different individuals i, or the model y_it = a_i + b t. Kreft and De Leeuw (1998) thus distinguish between fixed and random coefficients.

(2) Effects are fixed if they are interesting in themselves or random if there is interest in the underlying population. Searle, Casella, and McCulloch (1992, Section 1.4) explore this distinction in depth.

(3) "When a sample exhausts the population, the corresponding variable is *fixed*; when the sample is a small (i.e., negligible) part of the population the corresponding variable is *random*." (Green and Tukey, 1960)

(4) "If an effect is assumed to be a realized value of a random variable, it is called a random effect." (LaMotte, 1983)

(5) Fixed effects are estimated using least squares (or, more generally, maximum likelihood) and random effects are estimated with shrinkage ("linear unbiased prediction" in the terminology of Robinson, 1991). This definition is standard in the multilevel modeling literature (see, for example, Snijders and Bosker, 1999, Section 4.2) and in econometrics.

Gelman A. "ANOVA - why it is more important than ever - with Discussion", The Annals of Statistics, 2005, Vol. 33, No. 1, 1-53.

When should I treat a parameter as random?

Remember the adage: “All models are wrong, but some models are useful” -George Box

- Consider the model $Y_{ij} = \mu + \delta_i + \epsilon_{ij}$
- Suppose your goal was to estimate the average individual-specific effect $\mu = \frac{1}{r}\sum_{i=1}^{r} E(Y_{i1})$.
- What is the benefit of treating δ_i as a fixed effect vs. $\delta_i \sim (0, \sigma^2_\mu)$?

When should I treat a parameter as random?

Remember the adage: "All models are wrong, but some models are useful" -George Box

- Consider the model $Y_{ij} = \mu + \delta_i + \epsilon_{ij}$
- Suppose your goal was to estimate the average individual-specific effect $\mu = \frac{1}{r}\sum_{i=1}^{r} E(Y_{i1})$.
- What is the benefit of treating δ_i as a fixed effect vs. $\delta_i \sim (0, \sigma_\mu^2)$?

Properties of this random effect model

$$Y_{ij} = \mu + \delta_i + \epsilon_{ij} = \mu_i + \epsilon_{ij}$$

$$EY_{ij} = E(\mu_i) + E(\epsilon_{ij}) = \mu$$

$$\begin{aligned} Var(Y_{ij}) &= Var(\mu_i) + Var(\epsilon_{ij}) \\ &= \sigma_\mu^2 + \sigma^2 \end{aligned}$$

$$\begin{aligned} Cov(Y_{ij}, Y_{kl}) &= 0, \quad \text{for } i \neq k \\ Cov(Y_{ij}, Y_{ik}) &= Var(\mu_i) = \sigma_\mu^2, \text{for} \quad j \neq k \end{aligned}$$

Properties of this random effect model

$$Y_{ij} = \mu + \delta_i + \epsilon_{ij} = \mu_i + \epsilon_{ij}$$

$$EY_{ij} = E(\mu_i) + E(\epsilon_{ij}) = \mu$$

$$\begin{aligned} Var(Y_{ij}) &= Var(\mu_i) + Var(\epsilon_{ij}) \\ &= \sigma_\mu^2 + \sigma^2 \end{aligned}$$

$$\begin{aligned} Cov(Y_{ij}, Y_{kl}) &= 0, \quad \text{for } i \neq k \\ Cov(Y_{ij}, Y_{ik}) &= Var(\mu_i) = \sigma_\mu^2, \text{for} \quad j \neq k \end{aligned}$$

Properties of this random effect model

$$Y_{ij} = \mu + \delta_i + \epsilon_{ij} = \mu_i + \epsilon_{ij}$$

$$EY_{ij} = E(\mu_i) + E(\epsilon_{ij}) = \mu$$

$$\begin{aligned} Var(Y_{ij}) &= Var(\mu_i) + Var(\epsilon_{ij}) \\ &= \sigma_\mu^2 + \sigma^2 \end{aligned}$$

$$\begin{aligned} Cov(Y_{ij}, Y_{kl}) &= 0, \quad \text{for } i \neq k \\ Cov(Y_{ij}, Y_{ik}) &= Var(\mu_i) = \sigma_\mu^2, \text{for} \quad j \neq k \end{aligned}$$

Intraclass Correlation

- The intraclass correlation is the amount of variance accounted for by the random effect.
  - % of total variability attributed to the between subject variance
- The intraclass correlation is also the correlation between any two responses from the same subject.

$$
\begin{aligned}
Corr(Y_{ij}, Y_{ik}) &= \frac{Cov(Y_{ij}, Y_{ik})}{\sqrt{Var(Y_{ij})Var(Y_{ik})}} \\
&= \frac{\sigma_\mu^2}{\sigma^2 + \sigma_\mu^2} \\
&= \frac{Var(\mu_i)}{Var(Y_{ij})}
\end{aligned}
$$

Testing random effect

- Hypothesis of interest
  - $H_0 : \sigma_\mu^2 = 0$
  - vs $H_a : \sigma_\mu^2 > 0$
- Here $\sigma_\mu^2 = 0$ means
  - There is no subject effect.
  - All the variance in the observed data are due to random errors (i.i.d)
  - Two observations with in a subject are uncorrelated.

Approaches to testing H_0

$H_0 : \sigma_\mu^2 = 0.$

Sums-of-Squares/OLS approach:

- Can use various sums-of-squares to estimate variance components (σ^2 and σ_μ^2).
- Only "nice" if you have balanced data. Can easily be extend to unbalanced data.
- Will focus on sum-of-squares approach for balanced data today.

Maximum Likelihood Approach

- Write out the likelihood function for the model.
- Estimate variance components by maximizing the likelihood.
- Can be fit with `lme4` in R (or Proc Mixed in SAS).
- Valid for unbalanced data.
- We will rely on the assumption $\delta_i \sim N(0, \sigma_\mu^2)$ here (see next slide).

Assumptions when testing H_0

- $H_0 : \sigma^2_\mu = 0$
- Recall: for hypothesis testing of a fixed effect in the **mean model**, needed to get mean and variance correct. Why?
- In this case: parameter is σ^2_μ, which is a parameter in the **variance model**.
- Therefore need to get the variance (second moment) and variance of the variance (fourth moment) correct to perform accurate inference.
- We will rely on the assumption $\delta_i \sim N(0, \sigma^2_\mu)$, especially when using maximum likelihood.

Assumptions when testing H_0

- $H_0 : \sigma^2_\mu = 0$
- Recall: for hypothesis testing of a fixed effect in the **mean model**, needed to get mean and variance correct. Why?
- In this case: parameter is σ^2_μ, which is a parameter in the **variance model.**
- Therefore need to get the variance (second moment) and variance of the variance (fourth moment) correct to perform accurate inference.
- We will rely on the assumption $\delta_i \sim N(0, \sigma^2_\mu)$, especially when using maximum likelihood.

Assumptions when testing H_0

- $H_0 : \sigma_\mu^2 = 0$
- Recall: for hypothesis testing of a fixed effect in the **mean model**, needed to get mean and variance correct. Why?
- In this case: parameter is σ_μ^2, which is a parameter in the **variance model.**
- Therefore need to get the variance (second moment) and variance of the variance (fourth moment) correct to perform accurate inference.
- We will rely on the assumption $\delta_i \sim N(0, \sigma_\mu^2)$, especially when using maximum likelihood.

Here, we will use OLS. By doing so, we are implicitly modeling δ_i has a **fixed** effect.

- For simplicity: ASSUME BALANCED DATA: Y_{ij} where $i = 1, \ldots, r$ and $j = 1, \ldots, n$ (will extend on HW)
- Recall that in fixed one-way ANOVA, we have
  - $SSE = \sum_{ij}\left(Y_{ij} - \bar{Y}_{i.}\right)^2$ with df $r(n-1)$
  - $SSTR = n\sum_i\left(\bar{Y}_{i.} - \bar{Y}_{..}\right)^2$ with df $r-1$
  - $E(MSE) = \sigma^2$
  - $E(MSTR) = \sigma^2 + n\frac{\sum(\mu_i - \mu_.)^2}{r-1}$ where $\mu_. = \sum \mu_i / r$
- In a random effect model, we have the same SSE and $SSTR$, with
  - $E(MSE) = \sigma^2$, this is easy
  - For $MSTR$:

$$
\begin{aligned}
E\left(MSTR\right) &= E\left\{E\left(MSTR \mid \mu_1, \ldots, \mu_r\right)\right\} \\
&= \sigma^2 + E\left\{n\frac{\sum_i(\mu_i - \mu_.)^2}{r-1}\right\} = \sigma^2 + n\sigma_\mu^2
\end{aligned}
$$

Here, we will use OLS. By doing so, we are implicitly modeling δ_i has a **fixed** effect.

- For simplicity: ASSUME BALANCED DATA: Y_{ij} where $i = 1, \ldots, r$ and $j = 1, \ldots, n$ (will extend on HW)
- Recall that in fixed one-way ANOVA, we have
  - $SSE = \sum_{ij} \left(Y_{ij} - \bar{Y}_{i.}\right)^2$ with df $r(n-1)$
  - $SSTR = n \sum_i \left(\bar{Y}_{i.} - \bar{Y}_{..}\right)^2$ with df $r - 1$
  - $E(MSE) = \sigma^2$
  - $E(MSTR) = \sigma^2 + n\frac{\sum(\mu_i - \mu_.)^2}{r-1}$ where $\mu_. = \sum \mu_i / r$
- In a random effect model, we have the same SSE and $SSTR$, with
  - $E(MSE) = \sigma^2$, this is easy
  - For $MSTR$:

$$E\left(MSTR\right) = E\left\{E\left(MSTR \mid \mu_1, \ldots, \mu_r\right)\right\}$$
$$= \sigma^2 + E\left\{n\frac{\sum_i(\mu_i - \mu_.)^2}{r-1}\right\} = \sigma^2 + n\sigma_\mu^2$$

Here, we will use OLS. By doing so, we are implicitly modeling δ_i has a **fixed** effect.

- For simplicity: ASSUME BALANCED DATA: Y_{ij} where $i = 1, \ldots, r$ and $j = 1, \ldots, n$ (will extend on HW)
- Recall that in fixed one-way ANOVA, we have
  - $SSE = \sum_{ij} \left(Y_{ij} - \bar{Y}_{i.}\right)^2$ with df $r(n-1)$
  - $SSTR = n \sum_i \left(\bar{Y}_{i.} - \bar{Y}_{..}\right)^2$ with df $r - 1$
  - $E(MSE) = \sigma^2$
  - $E(MSTR) = \sigma^2 + n\frac{\sum(\mu_i - \mu_.)^2}{r-1}$ where $\mu_. = \sum \mu_i / r$
- In a random effect model, we have the same SSE and $SSTR$, with
  - $E(MSE) = \sigma^2$, this is easy
  - For $MSTR$:

$$E(MSTR) = E\{E(MSTR \mid \mu_1, \ldots, \mu_r)\}$$

$$= \sigma^2 + E\left\{n\frac{\sum_i(\mu_i - \mu_.)^2}{r-1}\right\} = \sigma^2 + n\sigma_\mu^2$$

SS and MS terms

Here, we will use OLS. By doing so, we are implicitly modeling δ_i has a **fixed** effect.

- For simplicity: ASSUME BALANCED DATA: Y_{ij} where $i = 1, \ldots, r$ and $j = 1, \ldots, n$ (will extend on HW)
- Recall that in fixed one-way ANOVA, we have
  - $SSE = \sum_{ij} \left(Y_{ij} - \bar{Y}_{i.}\right)^2$ with df $r(n-1)$
  - $SSTR = n \sum_i \left(\bar{Y}_{i.} - \bar{Y}_{..}\right)^2$ with df $r - 1$
  - $E(MSE) = \sigma^2$
  - $E(MSTR) = \sigma^2 + n\frac{\sum(\mu_i - \mu_.)^2}{r-1}$ where $\mu_. = \sum \mu_i / r$
- In a random effect model, we have the same SSE and $SSTR$, with
  - $E(MSE) = \sigma^2$, this is easy
  - For $MSTR$:

$$
\begin{aligned}
E\left(MSTR\right) &= E\left\{E\left(MSTR \mid \mu_1, \ldots, \mu_r\right)\right\} \\
&= \sigma^2 + E\left\{n\frac{\sum_i(\mu_i - \mu_.)^2}{r-1}\right\} = \sigma^2 + n\sigma_\mu^2
\end{aligned}
$$

Testing $H_0 : \sigma_\mu^2 = 0$

- Can show that $F^* = MSTR/MSE \sim F_{r-1,r(n-1)}$ when $\sigma_\mu^2 = 0$.
  - Stack to get $Y \sim N(\mu, V)$. V is NOT a multiple of the identity
  - $SSTR = n\sum_i \left(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}\right)^2 = Y'(H - \frac{1}{nr}11')Y$
  - Let $A = (H - \frac{1}{nr}11')/(n\sigma_\mu^2 + \sigma^2)$
  - $(r-1)MSTR/E(MSTR) = \frac{SSTR}{n\sigma_\mu^2+\sigma^2} = Y'AY$
  - AV is idempotent, $rank(A) = r - 1$
  - so $\frac{(r-1)MSTR}{n\sigma_\mu^2+\sigma^2} \sim \chi^2_{r-1}$.
  - Similarly, $\frac{r(n-1)MSE}{\sigma^2} \sim \chi^2_{r(n-1)}$.
  - can show that these are independent.
- You will show this for unbalanced designs on homework.

Distribution of Quad Form, χ^2 Distribution

- If $Y \sim N(\mu, V)$, then $Y^T AY \sim \chi^2_{rank(A),\mu^T A\mu}$, if and only if AV is idempotent.
- $rank(A)$ is the df, and $\mu^T A\mu$ is the non-central parameter.
- If $\mu^T A\mu = 0$, we call it χ^2 distribution, otherwise non-central χ^2 distribution.
- $\chi^2_{rank(A),\mu^T A\mu}$ has mean is $d + \lambda$ and variance is $2d + 4\lambda$

A quick note on some asymptotics

$$E(MSE) = \sigma^2, \quad E(MSTR) = \sigma^2 + n\sigma_\mu^2$$

- Consider the F statistic $F^* = MSTR/MSE$.
- If the alternative $H_A : \sigma_\mu^2 > 0$ is true, under what asymptotic regimes (as the sample size $nr \to \infty$) will your power to reject $H_0 : \sigma_\mu^2 = 0$ go to 1?
- Need $F^* \to \infty$ OR $MSTR \to E(MSTR) << \infty$ in probability to achieve perfect power.
- Expectation of denominator will be the same regardless of sample size. For simplicity, let's assume σ^2 is known.
- Expectation of numerator $\to \infty \Leftrightarrow n\sigma_\mu^2 \to \infty$. Is this realistic?
- Under usual assumptions, $MSTR \to E(MSTR)$ as $r \to \infty$. Is this realistic?

A quick note on some asymptotics

$$E(MSE) = \sigma^2, \quad E\left(MSTR\right) = \sigma^2 + n\sigma_\mu^2$$

- Consider the F statistic $F^* = MSTR/MSE$.
- If the alternative $H_A : \sigma_\mu^2 > 0$ is true, under what asymptotic regimes (as the sample size $nr \to \infty$) will your power to reject $H_0 : \sigma_\mu^2 = 0$ go to 1?
- Need $F^* \to \infty$ OR $MSTR \to E\left(MSTR\right) << \infty$ in probability to achieve perfect power.
- Expectation of denominator will be the same regardless of sample size. For simplicity, let's assume σ^2 is known.
- Expectation of numerator $\to \infty \Leftrightarrow n\sigma_\mu^2 \to \infty$. Is this realistic?
- Under usual assumptions, $MSTR \to E\left(MSTR\right)$ as $r \to \infty$. Is this realistic?

A quick note on some asymptotics

$$E(MSE) = \sigma^2, \quad E\left(MSTR\right) = \sigma^2 + n\sigma_\mu^2$$

- Consider the F statistic $F^* = MSTR/MSE$.
- If the alternative $H_A : \sigma_\mu^2 > 0$ is true, under what asymptotic regimes (as the sample size $nr \to \infty$) will your power to reject $H_0 : \sigma_\mu^2 = 0$ go to 1?
- Need $F^* \to \infty$ OR $MSTR \to E\left(MSTR\right) << \infty$ in probability to achieve perfect power.
- Expectation of denominator will be the same regardless of sample size. For simplicity, let's assume σ^2 is known.
- Expectation of numerator $\to \infty \Leftrightarrow n\sigma_\mu^2 \to \infty$. Is this realistic?
- Under usual assumptions, $MSTR \to E\left(MSTR\right)$ as $r \to \infty$. Is this realistic?

A quick note on some asymptotics

$$E(MSE) = \sigma^2, \quad E\,(MSTR) = \sigma^2 + n\sigma_\mu^2$$

- Consider the F statistic $F^* = MSTR/MSE$.
- If the alternative $H_A : \sigma_\mu^2 > 0$ is true, under what asymptotic regimes (as the sample size $nr \to \infty$) will your power to reject $H_0 : \sigma_\mu^2 = 0$ go to 1?
- Need $F^* \to \infty$ OR $MSTR \to E\,(MSTR) << \infty$ in probability to achieve perfect power.
- Expectation of denominator will be the same regardless of sample size. For simplicity, let's assume σ^2 is known.
- Expectation of numerator $\to \infty \Leftrightarrow n\sigma_\mu^2 \to \infty$. Is this realistic?
- Under usual assumptions, $MSTR \to E\,(MSTR)$ as $r \to \infty$. Is this realistic?

F Distribution

- Recall $F_{d1,d2}$: If $Z_1 \sim \chi^2_{d1}$ is independent of $Z_2 \sim \chi^2_{d2}$, then $\frac{Z_1/d1}{Z_2/d2} \sim F_{d1,d2}$
- Non-central F with degrees of freedom $d1$, $d2$ and non-centrality parameter λ generalize this:
  - $Z_1 \sim \chi^2_{d1,\lambda}$ is independent of $Z_2 \sim \chi^2_{d2}$
  - $\frac{Z_1/d1}{Z_2/d2} \sim F_{d1,d2,\lambda}$

- Adjustment can be made for unequal sample sizes.
  - Unbalanced data become a problem with more than one random effect.
  - We will focus on likelihood method for unbalanced data.

Estimation of ICC

- Remember the intraclass correlation coefficient
  - $ICC = \frac{\sigma_\mu^2}{\sigma^2+\sigma_\mu^2}$:= fraction of variance explained by individual
  - Correlation between two observations from the same subject.
- Some properties:
  - SSTR is independent of SSE
  - $SSTR \sim (\sigma^2 + n\sigma_\mu^2)\chi^2_{r-1}$
  - $SSE \sim \sigma^2\chi^2_{r(n-1)}$
- $\frac{MSTR}{n\sigma_\mu^2+\sigma^2}\frac{\sigma^2}{MSE} = \frac{MSTR}{MSE}\frac{\sigma^2}{\sigma^2+n\sigma_\mu^2} \sim F_{r-1,r(n-1)}$

- "Inverting" we can find that a $(1-\alpha)\%$ confidence interval for σ_μ^2/σ^2 is $[L, U]$.
  - $L = \frac{1}{n}\left[\frac{MSTR}{MSE}\left(\frac{1}{F_{r-1,r(n-1)}(1-\alpha/2)}\right) - 1\right]$
  - $U = \frac{1}{n}\left[\frac{MSTR}{MSE}\left(\frac{1}{F_{r-1,r(n-1)}(\alpha/2)}\right) - 1\right]$
- And $(1-\alpha)\%$ confidence interval for $ICC = \sigma_\mu^2/(\sigma^2+\sigma_\mu^2)$ is $[L^*, U^*]$.
  - $L^* = \frac{L}{1+L}$
  - $U^* = \frac{U}{1+U}$
- How to estimate *ICC*?
  1. Method of moments: $\frac{MSTR}{MSE}\frac{\sigma^2}{\sigma^2+n\sigma_\mu^2} \sim F_{r-1,r(n-1)} \Rightarrow$ $E(\frac{MSTR}{MSE}) = \frac{\sigma^2+n\sigma_\mu^2}{\sigma^2}\frac{r(n-1)-2}{r(n-1)}$
  2. Maximum likelihood.
- Can program yourself.

- "Inverting" we can find that a $(1-\alpha)\%$ confidence interval for σ_μ^2/σ^2 is $[L, U]$.
  - $L = \frac{1}{n}\left[\frac{MSTR}{MSE}\left(\frac{1}{F_{r-1,r(n-1)}(1-\alpha/2)}\right) - 1\right]$
  - $U = \frac{1}{n}\left[\frac{MSTR}{MSE}\left(\frac{1}{F_{r-1,r(n-1)}(\alpha/2)}\right) - 1\right]$
- And $(1-\alpha)\%$ confidence interval for $ICC = \sigma_\mu^2/(\sigma^2+\sigma_\mu^2)$ is $[L^*, U^*]$.
  - $L^* = \frac{L}{1+L}$
  - $U^* = \frac{U}{1+U}$
- How to estimate *ICC*?
  1. Method of moments: $\frac{MSTR}{MSE}\frac{\sigma^2}{\sigma^2+n\sigma_\mu^2} \sim F_{r-1,r(n-1)} \Rightarrow$
     $E(\frac{MSTR}{MSE}) = \frac{\sigma^2+n\sigma_\mu^2}{\sigma^2}\frac{r(n-1)-2}{r(n-1)}$
  2. Maximum likelihood.
- Can program yourself.

Estimation of σ^2 and σ_μ^2

- What if you wanted to estimate and get confidence intervals for σ^2 and σ_μ^2?
- σ^2 is easy.
  - MSE is an unbiased estimator and we know its distribution.
- σ_μ^2 is harder.
  - Do not have a mean sums of squares estimator of σ_μ^2.
  - Must take the linear combination $\frac{MSTR-MSE}{n}$.
  - Do not know the exact distribution of linear combinations of sums-of-squares.

Estimating σ^2

- Already showed that *MSE* is an unbiased estimator.
- $\frac{r(n-1)}{\sigma^2} MSE \sim \chi^2_{r(n-1)}$
- Invert the statistic to get $(1-\alpha)\%$ confidence interval:

$$\frac{r(n-1)MSE}{\chi^2_{r(n-1)}(1-\alpha/2)} \leq \sigma^2 \leq \frac{r(n-1)MSE}{\chi^2_{r(n-1)}(\alpha/2)}$$

Inference for σ_μ^2

- Note that $\sigma_\mu^2 = \frac{E(MSTR)}{n} - \frac{E(MSE)}{n}$.
- We do not know the distribution of linear combinations of sums-of-squares.
- There are several procedures for the approximation of this distribution:
  - We will focus on the Satterthwaite procedure.
  - Has some poor asymptotics with certain weights.
  - But you need a lot of subjects in general to get good variance component estimates.
- Consider $L = \sum_j c_j E(MS_j)$
  - MS_j is some mean square with degrees of freedom ν_j.
  - Unbiased estimator is $\hat{L} = \sum_j c_j MS_j$.

Satterthwaite

- Idea is to approximate $\frac{\nu \hat{L}}{L} \sim \chi^2_\nu$.
  - $\nu = \frac{\left(\sum_j c_j MS_j\right)^2}{\sum_j (c_j MS_j)^2/\nu_j}$
- CI: $\frac{\nu \hat{L}}{\chi^2_\nu(1-\alpha/2)} \leq L \leq \frac{\nu \hat{L}}{\chi^2_\nu(\alpha/2)}$
- To get a CI for σ^2_μ:
  - $MS_1 = MSTR$, $MS_2 = MSE$
  - $df_1 = r - 1$, $df_2 = r(n-1)$
  - $c_1 = n^{-1}$, $c_2 = -n^{-1}$
  - $\hat{L} = (MSTR - MSE)/n$
  - $df = \frac{n^2 \hat{L}^2}{MSTR^2/(r-1) + MSE^2/(r(n-1))}$
  - CI for σ^2_μ:
    $[\frac{\nu \hat{L}}{\chi^2_\nu(1-\alpha/2)}, \frac{\nu \hat{L}}{\chi^2_\nu(\alpha/2)}]$

Where does Satterthwaite come from?

A very useful tool: moment matching! Let $Z_j \sim \alpha_j \chi^2_{\nu_j}$ be independent, where α_j's are unknown, but ν_j's are known ($\alpha_j \propto \sigma_j^2$). We observe Z_j.

- $Z = \sum_j c_j Z_j$. We want to approximate $Z/\{E(Z)\}$ with $\nu^{-1}\chi^2_\nu$. Note $E\left(\nu^{-1}\chi^2_\nu\right) = 1$, Var $\left(\nu^{-1}\chi^2_\nu\right) = 2\nu^{-1}$
- First moment: $E\left[Z/\{E(Z)\}\right] = 1$, which matches.
- Variance (i.e. second moment):

$$\text{Var}\left[Z/\{E(Z)\}\right] = \frac{2}{E(Z)^2}\sum_j c_j^2 \alpha_j^2 \nu_j$$

  Gives us $\nu = \frac{E(Z)^2}{\sum_j c_j^2 \alpha_j^2 \nu_j}$. Problem: $E(Z), \alpha_j$ are unknown.
- Can replace $E(Z)$ with an unbiased estimator $Z = \sum_j c_j Z_j$.
- $E(Z_j) = \nu_j \alpha_j$. Replace $\alpha_j^2 \nu_j$ with Z_j^2/ν_j.

Where does Satterthwaite come from?

A very useful tool: moment matching! Let $Z_j \sim \alpha_j \chi^2_{\nu_j}$ be independent, where α_j's are unknown, but ν_j's are known ($\alpha_j \propto \sigma_j^2$). We observe Z_j.

- $Z = \sum_j c_j Z_j$. We want to approximate $Z/\{E(Z)\}$ with $\nu^{-1}\chi^2_\nu$. Note $E\left(\nu^{-1}\chi^2_\nu\right) = 1$, $\text{Var}\left(\nu^{-1}\chi^2_\nu\right) = 2\nu^{-1}$
- First moment: $E\left[Z/\{E(Z)\}\right] = 1$, which matches.
- Variance (i.e. second moment):

$$\text{Var}\left[Z/\{E(Z)\}\right] = \frac{2}{E(Z)^2}\sum_j c_j^2 \alpha_j^2 \nu_j$$

  Gives us $\nu = \frac{E(Z)^2}{\sum_j c_j^2 \alpha_j^2 \nu_j}$. Problem: $E(Z), \alpha_j$ are unknown.
- Can replace $E(Z)$ with an unbiased estimator $Z = \sum_j c_j Z_j$.
- $E(Z_j) = \nu_j \alpha_j$. Replace $\alpha_j^2 \nu_j$ with Z_j^2/ν_j.

Where does Satterthwaite come from?

A very useful tool: moment matching! Let $Z_j \sim \alpha_j \chi^2_{\nu_j}$ be independent, where α_j's are unknown, but ν_j's are known ($\alpha_j \propto \sigma_j^2$). We observe Z_j.

- $Z = \sum_j c_j Z_j$. We want to approximate $Z / \{E(Z)\}$ with $\nu^{-1}\chi^2_\nu$. Note $E\left(\nu^{-1}\chi^2_\nu\right) = 1$, $\mathrm{Var}\left(\nu^{-1}\chi^2_\nu\right) = 2\nu^{-1}$
- First moment: $E\left[Z / \{E(Z)\}\right] = 1$, which matches.
- Variance (i.e. second moment):

$$\mathrm{Var}\left[Z / \{E(Z)\}\right] = \frac{2}{E(Z)^2} \sum_j c_j^2 \alpha_j^2 \nu_j$$

Gives us $\nu = \frac{E(Z)^2}{\sum_j c_j^2 \alpha_j^2 \nu_j}$. Problem: $E(Z), \alpha_j$ are unknown.

- Can replace $E(Z)$ with an unbiased estimator $Z = \sum_j c_j Z_j$.
- $E(Z_j) = \nu_j \alpha_j$. Replace $\alpha_j^2 \nu_j$ with Z_j^2 / ν_j.

Where does Satterthwaite come from?

A very useful tool: moment matching! Let $Z_j \sim \alpha_j \chi^2_{\nu_j}$ be independent, where α_j's are unknown, but ν_j's are known ($\alpha_j \propto \sigma_j^2$). We observe Z_j.

- $Z = \sum_j c_j Z_j$. We want to approximate $Z/\{E(Z)\}$ with $\nu^{-1}\chi^2_\nu$. Note $E\left(\nu^{-1}\chi^2_\nu\right) = 1$, $\text{Var}\left(\nu^{-1}\chi^2_\nu\right) = 2\nu^{-1}$
- First moment: $E\left[Z/\{E(Z)\}\right] = 1$, which matches.
- Variance (i.e. second moment):

$$\text{Var}\left[Z/\{E(Z)\}\right] = \frac{2}{E(Z)^2}\sum_j c_j^2\alpha_j^2\nu_j$$

  Gives us $\nu = \frac{E(Z)^2}{\sum\limits_j c_j^2\alpha_j^2\nu_j}$. Problem: $E(Z), \alpha_j$ are unknown.

- Can replace $E(Z)$ with an unbiased estimator $Z = \sum_j c_j Z_j$.
- $E(Z_j) = \nu_j\alpha_j$. Replace $\alpha_j^2\nu_j$ with Z_j^2/ν_j.