SOLUTIONS TO HOMEWORK 4

1.1 The general rule of thumb is that if the experiment were to be repeated, any factors whose levels would be identical to the current experiment are fixed. Those whose levels would change randomly are random. From this rule, Factors C, B and Y are all fixed factors because the companies, brands of machines and types of yarn would be identical. Note that C, B and Y are crossed factors. Factor M is random because the machines in the experiment are selected at random from the 20 machines of a given brand. Note that M is nested within C$*$B. Worker (not specifically identified as a factor) is also a random factor, since the workers in the experiment are sampled (at random, we assume) from the work force at each of the companies. The differences between replications for a given worker operating a given machine and using a given type of yarn is the source of the error term in the model.

2.1 Let U and V be independent χ^2 random variables with $m-1$ and $N-m$ degrees of freedom, respectively, and let

$$W_N = N\log\left(1+\frac{U}{V}\right).$$

Note that V is distributed as $\sum_{i=1}^{N-m} Z_i^2$, so that, according to the Law of Large Numbers, $V/(N-m) \to 1$ in probability as $N \to \infty$. Therefore $U/V \to 0$ in probability.

From the inequality $x - x^2/2 \le \log(1+x) \le x$, valid if $x \ge 0$, we have

$$N\left[\frac{U}{V} - \frac{1}{2}\left(\frac{U}{V}\right)^2\right] \le W_N \le N\left(\frac{U}{V}\right).$$

We can write

$$\frac{NU}{V} = \frac{N}{N-m}\frac{U}{V/(N-m)} \quad \text{and} \quad \frac{NU^2}{V^2} = \frac{N}{(N-m)^2}\left(\frac{U}{V/(N-m)}\right)^2.$$

Therefore, $NU/V \to U$ and $NU^2/V^2 \to 0$ in probability, which means that $W_N \to U$ in probability. Moreover, the system of inequalities above implies that $W_N \to U$ in probability and hence in distribution.

Let $q = \chi_{m-1,1-\alpha}$ denote the $1-\alpha$ quantile of the χ^2 distribution with $m-1$ degrees of freedom and let G denote the χ^2_{m-1} cdf. In addition, let

$$w_N = N\log\left(1+\frac{m-1}{N-m}F_{m-1,N-m,1-\alpha}\right)$$

where $F_{m-1,N-m,1-\alpha}$ is the $1-\alpha$ quantile of the F distribution with $(m-1, N-m)$ degrees of freedom. We must show that $w_N \to q$.

Let $\varepsilon > 0$ be given. Using the continuity of G, choose δ so small that

$$G(q-\varepsilon) < 1-\alpha-\delta < 1-\alpha < 1-\alpha+\delta < G(q-\varepsilon)$$

Then choose N_0 so large that if $N > N_0$,

$$|P[W_N \le q-\varepsilon] - G(q-\varepsilon)| < \delta/2 \quad \text{and} \quad |P[W_N \le q-\varepsilon] - G(q-\varepsilon)| < \delta/2$$

Then by combining these inequalities we have

$$P[W_N \le q-\varepsilon] < 1-\alpha-\delta/2 < 1-\alpha < 1-\alpha+\delta/2 < P[W_N \le q-\varepsilon].$$

But $1-\alpha = P[W_N \le w_N]$ by definition, and by the monotonicity of cdf's, when $N > N_0$, $q-\varepsilon < w_N < q+\varepsilon$. Since ε was arbitrary, $w_N \to q$ and the proof is complete.

2.6 Write

$$\begin{aligned}\log L &= -\frac{N-m}{2}\log\sigma_e^2 - \frac{1}{2}\sum_{i=1}^m \log\lambda_i - \frac{SSE}{2\sigma_e^2} - \sum_{i=1}^m \frac{n_i(\bar{Y}_{i\cdot}-\mu)^2}{2\lambda_i} \\ &= -\frac{N-m}{2}\log\sigma_e^2 - \frac{1}{2}\sum_{i=1}^m \log(n_i\sigma_a^2+\sigma_e^2) - \frac{SSE}{2\sigma_e^2} - \sum_{i=1}^m \frac{n_i(\bar{Y}_{i\cdot}-\mu)^2}{2(n_i\sigma_a^2+\sigma_e^2)}\end{aligned}$$

where $\lambda_i = n_i\sigma_a^2 + \sigma_e^2$ in the notation of the text.

As $\sigma_e^2 \to 0$, observe that $\sigma_e^2 \log\sigma_e^2 \to 0$. Therefore $\log L$ can be written as

$$-\frac{1}{\sigma_e^2}\left[\frac{N-m}{2}\sigma_e^2\log\sigma_e^2 + \frac{SSE}{2}\right] - \frac{1}{2}\sum_{i=1}^m \log(n_i\sigma_a^2+\sigma_e^2) - \sum_{i=1}^m \frac{n_i(\bar{Y}_{i\cdot}-\mu)^2}{2(n_i\sigma_a^2+\sigma_e^2)}$$

which tends to $-\infty$ as $\sigma_e^2 \to 0$.

As $\sigma_e^2 \to \infty$, observe that the logarithmic terms in $\log L$ become infinite and the rational terms tend to zero. Since all terms in $\log L$ become negative, clearly $\log L \to -\infty$ as $\sigma_e^2 \to \infty$.

2.8 Rewrite

$$T = \frac{\bar{Y}_{\cdot\cdot} - \mu}{\sqrt{\mathrm{MSA}/(mn)}} = \frac{(\bar{Y}_{\cdot\cdot}-\mu)/\sqrt{(\sigma_a^2+\sigma_e^2/n)/m}}{\sqrt{\mathrm{MSA}/(m\sigma_a^2+\sigma_e^2)}}.$$

From the model $Y_{ij} = \mu + a_i + e_{ij}$ we see that $\bar{Y}_{\cdot\cdot} = (1/m)\sum_i(\mu + a_i + \bar{e}_{i\cdot})$, which has a $N(\mu, (\sigma_a^2 + \sigma_e^2/n)/m)$ distribution. Therefore the numerator in the rightmost displayed expression is standard normal.

From the formula

$$\mathrm{SSA} = \sum_i n[a_i + \bar{e}_{i\cdot} - (bara. - \bar{e}_{\cdot\cdot})^2$$

we can see that SSA is a sum of squared deviations of independent normal variables and is therefore independent of $\bar{Y}_{\cdot\cdot}$ and is distributed as $(\sigma_a^2 + \sigma_e^2/n)\chi^2$ with $m-1$ d.f. Therefore T has a Student t distribution with $m-1$ d.f.

6.5 The statistic MSA/MSE is distributed like

$$\left(1+\frac{m\sigma_a^2}{\sigma_e^2}\right)\mathcal{F}=(1+m\theta)\mathcal{F}$$

where $\mathcal{F}$ has an F distribution with $m-1$ and $m(n-1)$ degrees of freedom. Let $F_1 < F_2$ denote the lower and upper $\alpha/2$ percentage points of this distribution. Note that $\rho=\sigma_a^2/(\sigma_a^2+\sigma_e^2)=\theta/(1+\theta)$. Knowing the distribution of MSA/MSE, we can derive confidence limits for θ as follows:

$$\begin{aligned}
1-\alpha &= P\left[\left(1+\frac{m\sigma_a^2}{\sigma_e^2}\right)F_1<\frac{\mathrm{MSA}}{\mathrm{MSE}}<\left(1+\frac{m\sigma_a^2}{\sigma_e^2}\right)F_2\right] \\
&= P\left[\frac{1}{m}\left(\frac{\mathrm{MSA}}{F_2\mathrm{MSE}}-1\right)<\theta<\frac{1}{m}\left(\frac{\mathrm{MSA}}{F_1\mathrm{MSE}}-1\right)\right] \\
&= P[\hat{\theta}_L<\theta<\hat{\theta}_U].
\end{aligned}$$

The required confidence limits for ρ come from transforming the limits for θ, so that $\hat{\rho}_L=\hat{\theta}_L/(1+\hat{\theta}_L)$ and $\hat{\rho}_U=\hat{\theta}_U/(1+\hat{\theta}_U)$.

6.7 The "usual" SSA can be rewritten SSA $=\sum_{i=1}^m n_i\bar{Y}_{i\cdot}^2-N\bar{Y}_{\cdot\cdot}$ and the unweighted sum of squares is SSU $=\sum_{i=1}^m \bar{Y}_{i\cdot}^2-m\bar{Y}_u^2$, where

$$\bar{Y}_{\cdot\cdot}=\frac{1}{N}\sum_{i=1}^m\sum_{j=1}^{n_i}Y_{ij}=\frac{1}{N}\sum_{i=1}^m n_i\bar{Y}_{i\cdot}\quad\text{and}\quad \bar{Y}_u=\frac{1}{m}\sum_{i=1}^m\bar{Y}_{i\cdot}$$

are the weighted and unweighted means of sample group averages. Under the model, $\bar{Y}_{i\cdot}\sim N(\mu,\sigma_a^2+\sigma_e^2/n_i)$, so we have $E[\bar{Y}_{i\cdot}^2]=\mu^2+\sigma_a^2+\sigma_e^2/n_i$. Therefore

$$\begin{aligned}
E[\mathrm{SSA}] &= (m-1)\left[N\mu^2+\sum_{i=1}^m(n_i\sigma_a^2+\sigma_e^2)-N\mu^2-\sum_{i=1}^m n_i^2(\sigma_a^2+\sigma_e^2/n_i)/N^2\right] \\
&= \left[N-\frac{1}{N}\sum_{i=1}^m n_i^2\right]\sigma_a^2+(m-1)\sigma_e^2
\end{aligned}$$

Since MSE is an unbiased estimator of σ_e^2, we see that

$$\frac{\mathrm{SSA}-(m-1)\mathrm{MSE}}{[N-(1/N)\sum_{i=1}^m n_i^2]}$$

is an unbiased estimator of σ_a^2. Similarly one can show

$$\begin{aligned} E[\mathrm{SSU}] &= \left[m\mu^2 + m\sigma_a^2 + \sum_{i=1}^{m} \sigma_e^2/n_i - m\mu^2 - \sigma_a^2 - \sum_{i=1}^{m} \sigma_e^2/(mn_i)\right] \\ &= (m-1)\sigma_a^2 + (m-1)\sigma_e^2 \sum_{i=1}^{m} 1/(mn_i) \end{aligned}$$

Therefore $\mathrm{SSU}/(m-1) - \mathrm{MSE} \sum_{i=1}^{m} 1/(mn_i)$ is also an unbiased estimator of σ_a^2. Both estimators depend on the sufficient statistics for the problem, namely $\bar{Y}_{i\cdot}$, $i = 1, \ldots, m$ and SSE and are clearly different.