1(a) Since $Y_{ij} = M_i + E_{ij}$ where $E_{ij} \sim iid N(0, \sigma^2)$, $Y_{ij}$'s are independently distributed and $Y_{ij} \sim N(M_i, \sigma^2)$.

So the likelihood function is

$$L(M_1, \ldots, M_k, \sigma^2) = \prod_{i=1}^{k} f(Y_{ij} \mid M_{i}, \sigma^2)$$

$$= \prod_{i=1}^{k} \prod_{j=1}^{J} f(Y_{ij} \mid M_{i}, \sigma^2)$$

$$= \prod_{i=1}^{k} \left[ \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(Y_{ij} - M_i)^2}{2\sigma^2} \right\} \right]$$

Because the log function is monotonically increasing, maximizing $L$ is equivalent to maximizing

$$l = \log(L) = \sum_{i=1}^{k} \sum_{j=1}^{J} \log \left[ \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(Y_{ij} - M_i)^2}{2\sigma^2} \right\} \right]$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{J} \left[ -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{(Y_{ij} - M_i)^2}{2\sigma^2} \right]$$

$$= - \frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{\sum_{i=1}^{k} \sum_{j=1}^{J} (Y_{ij} - M_i)^2}{2\sigma^2}$$

Because of the "-" in $l$, maximizing $L$ is equivalent to minimizing $\sum_{i=1}^{k} \sum_{j=1}^{J} (Y_{ij} - M_i)^2$.
(b) To obtain MLE, we take partial derivative, set it to 0, and solve it:

\[
\frac{\partial \theta}{\partial u_i} = \frac{2}{3} u_i \left( \sum_j (x_{ij} - u_i)^2 \right) = \frac{2}{3} u_i \left( \sum_j (x_{ij} - u_i)^2 + \sum_{i \neq i} (x_{ij} - u_i)^2 \right) = \sum_j 2(x_{ij} - u_i)(-1) + 0
\]

\[= -2 \sum_j (x_{ij} - J u_i)\]

Set it to 0, we have

\[\sum_j \frac{x_{ij}}{J} = J u_i\]

So \( \hat{u}_i = \frac{\sum_j x_{ij}}{J} = \bar{x}_i \).

(c) \( \{x_{ij}, \ldots, y_{ij}\} \) is a random sample from \( N(u_i, \sigma^2) \), as \( x_{ij} \) and \( y_{ij} \) are independent and have the same distribution. 

\[\bar{x}_i = \frac{1}{J} \sum_j x_{ij}\] is the sample mean. By what we learned in 120B, \( \bar{x}_i \) follows a normal distribution. Since \( \mathbb{E}[\bar{x}_i] = \frac{1}{J} \sum_j \mathbb{E}[x_{ij}] \)

\[\text{Var}[\bar{x}_i] = \left( \frac{1}{J} \right)^2 \text{Var}[x_{ij}] = \frac{1}{J^2} J \sigma^2 = \frac{\sigma^2}{J}, \quad \bar{x}_i \sim N(u_i, \frac{\sigma^2}{J}) \]
The $\mu_i$'s are independent, because they are from independent samples. When $\mu_1 = \ldots = \mu_i = \mu_i$, they all follow $N(\mu, \sigma^2)$. So when $\mu_1 = \ldots = \mu_i$, $[\bar{Y}_1, \ldots, \bar{Y}_i]$ is a random sample from $N(\mu_i, \sigma^2/\sqrt{i})$. Let $S^2$ denote the sample variance, i.e.,

$$S^2 = \frac{1}{i-1} \sum \left( \bar{Y}_i - \mu_i \right)^2$$

By 120 B, $\frac{(i-1)S^2}{\sigma^2/\sqrt{i}} \sim \chi^2_{i-1}$

But $\frac{(i-1)S^2}{\sigma^2/\sqrt{i}} = \frac{i \sum (\bar{Y}_i - \mu_i)^2}{\sigma^2} = \frac{SS_{i} B}{\sigma^2}$

So $\frac{SS_{i} B}{\sigma^2} \sim \chi^2_{i-1}$.

(a) Consider $\bar{Y}_1, \ldots, \bar{Y}_i$. It is a random sample from $N(\mu_i - \mu, \sigma^2)$. By 120 B,

$$\frac{(i-1)\hat{S}_i^2}{\sigma^2} = \frac{\sum (\bar{Y}_j - \mu)^2}{\sigma^2} \sim \chi^2_{i-1}$$

where $\hat{S}_i^2$ is the sample variance.

$\hat{S}_1^2, \ldots, \hat{S}_i^2$ are independent, because they are from independent samples. As a result,

$$\frac{i \sum (\bar{Y}_j - \mu)^2}{\sigma^2} \sim \chi^2_{i-1} \Rightarrow \chi^2_{i-1}$$
But \[ \frac{\Sigma (y_{ij} - \bar{y}_{..})^2}{\sigma^2} \leq \frac{\Sigma (y_{ij} - \bar{y}_{..})^2}{\sigma^2} = \frac{SSW}{\sigma^2} \]

So \[ \frac{SSW}{\sigma^2} \sim \chi^2_{(I-1)J} \]

(e) No. To make a F-statistic we need independence, i.e. SSB \ perpendicular \ SSW.

(f) \[ F = \frac{SSB/(I-1)}{SSW/[I(J-1)]} \sim F_{I-1, I(J-1)} \]

2. (a) \( I = 3, J = 3, \ k = 6 \)

(b) \( \epsilon_{ijk} \sim N(0, \sigma^2) \)

(c) \( \mu \) : the overall mean

\( \alpha_i \) : the main effect of technician \( i \)

\( \beta_j \) : the main effect of disk type \( j \)

\( \alpha \beta_{ij} \) : the interaction effect between technician \( i \) and disk type \( j \).

(d) \( SSTO = SSA + SSB + SSAB + SSE \)

where \[ SSA = \frac{1}{jk} \sum_{i,j,k} (\bar{y}_{...} - \bar{y}_{..})^2 = \frac{1}{jk} \sum_{i,j,k} (\bar{y}_{...} - \bar{y}_{..})^2 \]

\[ SSB = \frac{1}{ik} \sum_{j,k} (\bar{y}_{j...} - \bar{y}_{..})^2 = \frac{1}{ik} \sum_{j,k} (\bar{y}_{j...} - \bar{y}_{..})^2 \]

\[ SSAB = \frac{1}{jk} \sum_{i,j,k} (\bar{y}_{ij.} - \bar{y}_{..} + \bar{y}_{..} - \bar{y}_{...})^2 = \frac{1}{jk} \sum_{i,j,k} (\bar{y}_{ij.} - \bar{y}_{..} + \bar{y}_{..} - \bar{y}_{...})^2 \]

\[ SSE = \frac{1}{ijk} \sum_{i,j,k} (y_{ijk} - \bar{y}_{ij.} - \bar{y}_{j.} + \bar{y}_{..})^2 \]
(c) Source df SS MS F
   type  2  100  50  7.5
   tech  2  10  5  0.75
   Interaction  4  90  22.5  3.375
   Error  45  300  20/3  —
   Total  53  500  —  —

An unbiased estimate is \( \text{MSE} = \frac{20}{3} \)

(g) From the table on page 1 of the exam

\[ F_{2, 45, 0.95} = 3.204 \]

Since \( F = 7.5 > F_{2, 45, 0.95} \), we reject the null hypothesis and conclude that type has an effect on service time.

(h) The null hypothesis is

\( H_0: \) the effect of type on service time does not depend on technician

(or the effect of technician on service time does not depend on type)

\( H_1: \) the effect of type on service time depends on technician

(or the effect of technician on service time depends on type)
Alternatively, you can also write:

\[ H_0: \beta_1 = \ldots = \beta_3 = 0 \]
\[ H_A: \text{not all } \beta_j \text{'s equal } 0 \]

The F-statistic for this problem is

\[ F = \frac{SSAB/4}{SSE/45} = 3.375 \]

From the table on page 1 of the exam

\[ F_{4.45, .05} = 2.579 \]

Since \( F = 3.375 > F_{4.45, .05} \), we reject \( H_0 \). So we conclude that the effect of type on service time depends on technician.

3. (a) Do female students have higher GPA than male students?
(b) Do human's left feet longer than right feet?