Sol:
(a).
The Splus procedure is as follows:

```r
> rad=read.table("h:\stat524\Datasets\mineral.data",header=T)
> rad
   Cr  Sr
  1 0.48 12.57
  2 40.53 73.68
  3  2.19 11.13
  4  0.55 20.03
  5  0.74 20.29
  6  0.66  0.78
  7  0.93  4.64
  8  0.37  0.43
  9  0.22  1.08

> xbar=apply(rad,2,mean)
> xbar
   Cr  Sr
  5.185556 16.07

> S=var(rad)

> n=nrow(rad)
> p=ncol(rad)
> fcrit=((n-1)*p/(n-p))*qf(0.9,p,n-p)
> plot(ellipse(S/42,centre=xbar,t=sqrt(fcrit)),type="l")
```

And we got the picture as:
(b).
The simultaneous 90% CI is calculated as follows:

\[
\begin{align*}
&s_{intv1} = c(xbar[1] - \text{sqrt}(S[1,1]*fcrit/n), \text{xbar}[1]+\text{sqrt}(S[1,1]*fcrit/n)) \\
&s_{intv2} = c(xbar[2] - \text{sqrt}(S[2,2]*fcrit/n), \text{xbar}[2]+\text{sqrt}(S[2,2]*fcrit/n))
\end{align*}
\]

Thus it does seem that a mean strontium level of 10 drops into this CI.

In fact, since Cr and Sr cannot assume negative numbers, we can reduce the CI to:

\[
\begin{align*}
\text{Cr} & \quad (0, 17.25228) \\
\text{Sr} & \quad (0, 36.96696)
\end{align*}
\]

Therefore, if we consider the 2 means separately, we can say that our data doesn't reject either of the 2 means (0.3, or 10) from 90% confidence.

\[
\text{Also the point (0.3,10) falls inside the 90% confidence ellipsoid, whence the value (0.3, 10) is a plausible value for the mean vector.}
\]
\( t^2 = 1.772481 < f_{crit} = 7.445582 \). Thus we can also conclude that we don’t reject the hypothesis of \((0.3, 10)'\).

(c).
The QQ plot of \( Cr \) is

![QQ plot of Cr](image1)

The QQ plot of \( Sr \) is:

![QQ plot of Sr](image2)

The scatter diagram is
Since none of the 2 QQ plots shows any linearity property, and neither does the scatter diagram show any normality, we doubt the fact of normality. Thus our conclusions from (a) and (b) are not valid.

(d).
The following 2 plots are the QQ plots for the data set without the obvious outlier.
From the plots, we still doubt the normality. But we still calculate the inferences.

```r
> xbar_apply(rad,2,mean)
> xbar
Cr  Sr
0.7675 8.86875
> S_var(rad)
> S
Cr  Sr
Cr 0.3785643 1.030282
Sr 1.0302821 69.859755
> sintv1_c(xbar[1]-sqrt(S[1,1]*fcrit/n), xbar[1]+sqrt(S[1,1]*fcrit/n))
> sintv2_c(xbar[2]-sqrt(S[2,2]*fcrit/n), xbar[2]+sqrt(S[2,2]*fcrit/n))
> sintv1
Cr  Cr
0.1491156 1.385884
> sintv2
Sr  Sr
0.468306 17.26919
```

From the following, we claim that (0.3, 10) is not rejected.

```r
> mu_c(0.3,10)
> t2_3 (xbar-mu) %*% solve(S) %*% (xbar-mu)
> t2
[,1]
[1,] 5.30
> fcrit
[1] 8.081043
>
(a) Construct and sketch a 95% confidence ellipse for the pair \([\mu_1, \mu_2]^T\), where \(\mu_1 = E[X_1]\) and \(\mu_2 = E[X_2]\).

Ans: First use EM algorithm to find the mean and variance-covariance matrix

\[
\mu = \left( \frac{\sum_{i \neq 24} x_{i,1} + x_{24,1}}{\sum_{i \neq 24} x_{i,1} + x_{24,1}} \right), \quad \Sigma = \left( \begin{array}{cc} \frac{x_{24,1}^2 + \sum_{i \neq 24} x_{i,1}^2}{x_{24,1}^2 + x_{24,1}x_{i,1}^2 + x_{i,1}x_{24,1}} & \frac{x_{24,1}x_{i,1}}{x_{24,1}^2 + x_{24,1}x_{i,1}^2 + x_{i,1}x_{24,1}} \\ \frac{x_{i,1}x_{24,1}}{x_{i,1}x_{24,1} + x_{24,1}x_{i,1}^2 + x_{i,1}^2} & \frac{x_{i,1}^2}{x_{i,1}x_{24,1} + x_{24,1}x_{i,1}^2 + x_{i,1}^2} \end{array} \right)
\]

Below is the Splus program for this problem:

```splus
ft2_function(x) {
  T2 <- matrix(0,2,2)
  n <- length(x[,1])
  for (i in 1:n) {
    T2[1,1] <- T2[1,1] + (x[i,1])^2
    T2[2,2] <- T2[2,2] + (x[i,2])^2
    T2[1,2] <- T2[1,2] + (x[i,1]*x[i,2])
    T2[2,1] <- T2[2,1] + (x[i,2]*x[i,1])
  }
  return(T2)
}

sigma <- function(mu, T2, n) {
  S <- T2/n - (mu)%*%t(mu)
  return(S)
}

updatet2_function(x, mu, sigma) {
  n <- length(x[,1])
  new8.2 <- mu[2] + sigma[2,1]/sigma[1,1] * (x[8,1] - mu[1])
  new8.2.sqr <- sigma[2,2] - sigma[2,1]/sigma[1,1] * sigma[1,2] + (new8.2)^2
  new24.1.sqr <- sigma[1,1] - sigma[1,2]/sigma[2,2] * sigma[2,1] + (new24.1)^2
  new24.1t2 <- new24.1 * x[24,2]
  new8.1t2 <- new8.2 * x[8,1]
  S <- matrix(0,2,2)
  S[1,1] <- S[1,1] + new24.1.sqr
  S[1,2] <- S[1,2] + new8.1t2 + new24.1t2
  for (i in 1:n) {
    if (i!=24) {
      S[1,1] <- S[1,1] + (x[i,1])^2
    }
    if (i!=8) {
      S[2,2] <- S[2,2] + (x[i,2])^2
    }
    if ((i!=8) && (i!=24)) {
      S[1,2] <- S[1,2] + x[i,1]*x[i,2]
      S[2,1] <- S[2,1] + x[i,2]*x[i,1]
    }
  }
  return(S)
}

p2_function() {

```
```r
lumber_read.table("e:\stat524\Dataset\lumber.data", header=T)
n_length(lumber[,1])
lumber[8,2] = 0
lumber[24,1] = 0

mu_matrix((apply(lumber,2,mean)*30/29),2,1)
lumber[24,1] = mu[1]

T2_ft2(lumber)
sigma0_sigma(mu,T2,n)
print(0)
print(mu)
print(sigma0)

for (i in 1:6) {
lumber[8,2] = mu[2] + sigma0[2,1]/sigma0[1,1]*(lumber[8,1] - mu[1])
lumber[24,1] = mu[1] + sigma0[1,2]/sigma0[2,2]*(lumber[24,2] - mu[2])

mu_matrix(apply(lumber,2,mean),2,1)
T2_updateT2(lumber,mu,sigma0)
sigma0_sigma(mu,T2,n)
print(i)
print(mu)
print(sigma0)
}
}

After 6 iteration, the \( \hat{\mu} \) and \( \hat{\Sigma} \) converge.

> p2()
[1] 0
 [,1] 1869.931
[2,] 8353.759
 [,1] 117340.1
[2,] 349165.5
[1] 1
 [,1] 1869.776
[2,] 8347.219
 [,1] 120111.3
[2,] 349381.8
[1] 2
 [,1] 1869.797
[2,] 8347.163
 [,1] 120166.7
[2,] 349605.7
[1] 3
 [,1] 1869.798
[2,] 8347.167
```

6
The estimated mean and variance-covariance matrix are:

\[ \hat{\mu} = \begin{pmatrix} 1869.798 \\ 8347.158 \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} 120166.1 & 349605.8 \\ 349605.8 & 3452562.8 \end{pmatrix} \]

Using this result, the 95% confidence ellipse is showing below.

(b) Suppose \( \mu_10 = 2000 \) and \( \mu_20 = 10,000 \) represent ‘typical’ values for stiffness and bending strength, respectively. Given the result in (a), are the data in table consistent with these values? Explain.
Ans: The 95% simultaneous confidence interval for $\mu_1$ is (1703.317, 2036.279) and for $\mu_2$ is (7454.791, 9239.525)

```r
sintv11 <- c(xbar[1]-sqrt(xvar[1,1]*fcrit/n), xbar[1]+sqrt(xvar[1,1]*fcrit/n))
> sintv11
[1] 1703.317 2036.279
> sintv12
[1] 7454.791 9239.525
```

2000 is inside the simultaneous CI of $\mu_1$, but 10000 is outside the simultaneous CI of $\mu_2$. At significant level $\alpha = 0.05$, the data is not consistent with the hypothesis of \[
\begin{pmatrix}
\mu_1 \\
\mu_2
\end{pmatrix}
= \begin{pmatrix}
2000 \\
10000
\end{pmatrix}.
\]

(c) Is the bivariate normal distribution a viable population model? Explain with reference to Q-Q plots and a scatter diagram.

Ans: The QQ-plot and scatter plot are showing below. From the plot, I conclude that the data follows bivariate normal distribution.