PROBABILISTIC METHODS IN ENGINEERING

VE401

Assignment III

Due: March 29, 2018

Team number: 20

Team members:

Jiecheng SHI	515370910022
Shihan ZHAN	516370910128
Tianyi GE	516370910168

Instructor:

Prof. Horst HOHBERGER

May 15, 2018

Exercise 3.1

1) We know that

$$\begin{aligned}
&P[\text{a random sample of 50 pages contains at least one error}]\\
=&1-P[\text{a random sample of 50 pages contains no error}]\\
=&1-(\frac{150}{200})^5\\
\approx&0.763.
\end{aligned}$$

2) We denote that the probability of being sampled is p, then

$$\begin{aligned}
&P[\text{a random sample of 50 pages contains at least three error}]\\
=&\sum_{k=3}^{5}\binom{5}{k}p^k(1-p)^{5-k}\\
\approx&1-\Phi(\frac{3-1/2-5p}{\sqrt{5p(1-p)}})\\
\overset{!}{=}&90\%.
\end{aligned}$$

Thus

$$\begin{aligned}
&\frac{3-1/2-5p}{\sqrt{5p(1-p)}}=-1.29,\\
&p=0.750.\\
&n=200p\approx150.
\end{aligned}$$

Hence, the random sample must contain 150 pages.

Exercise 3.2

1)

$$\begin{aligned}
R_1(t)&=1-\int_0^t 0.003x^{-0.5}e^{-0.006x^{0.5}}\mathrm{d}x=1-(-e^{-0.006x^{0.5}})\Big|_0^t=e^{-0.006t^{0.5}}.\\
R_2(t)&=1-\int_0^t\frac{1}{25000}e^{-\frac{x}{25000}}=1-(-e^{-\frac{x}{25000}}\Big|_0^t)=e^{-t/25000}.\\
R_s(t)&=\prod_{i=1}^{2}R_1(t)R_2(t)=e^{-0.006t^{0.5}-t/25000}.
\end{aligned}$$

Hence, at 2500 hours,

$$R_s(2500)\approx0.67.$$

2)

$$F_s(t) = 1 - R_s(t) = 1 - e^{-0.006t^{0.5} - t/25000}.$$

When $t < 2000$, the probability that the system fails is

$$F_s(2000) = 0.29.$$

3)

$$\begin{aligned} R_p(t) = 1 - \prod_{i=1}^{2}(1 - R_1(t))(1 - R_2(t)) \\ =& R_1(t) + R_2(t) - R_1(t)R_2(t) \\ =& e^{-0.006t^{0.5}} + e^{-t/25000} - e^{-0.006t^{0.5} - t/25000}. \end{aligned}$$

Hence, at 2500 hours,

$$R_p(t) \approx 0.98.$$

Exercise 3.3

1)

$$\sum_{x=1}^{n}\sum_{y=x}^{n} f_{XY}(x,y) = \sum_{x=1}^{n}\sum_{y=x}^{n} \frac{2}{n(n+1)} = \frac{n(n+1)}{2}\frac{2}{n(n+1)} = 1.$$

Also, $f_{XY}(x,y) >= 0$. Hence, $f_{XY}(x,y)$ is a density.

2) To find the marginal density,

$$f_X(x) = \sum_{y=x}^{n} \frac{2}{n(n+1)} = \frac{2(n-x+1)}{n(n+1)}.$$

$$f_Y(y) = \sum_{x=1}^{y} \frac{2}{n(n+1)} = \frac{2y}{n(n+1)}.$$

3)

$$\begin{aligned} f_X(x)f_Y(y) &= \frac{2(n-x+1)}{n(n+1)}\frac{2y}{n(n+1)} \\ &= \frac{4y(n-x+1)}{n^2(n+1)^2} \neq \frac{2}{n(n+1)}. \end{aligned}$$

Thus, X and Y are not independent.

4)

$$P[X \leq 3 \text{ and } Y \leq 2] = \sum_{x=1}^{3} \sum_{y=x}^{2} \frac{2}{5 \cdot (5+1)} = \frac{1}{15} \times 3 = \frac{1}{5}.$$

Exercise 3.4

We denote that $V = Y$. Considering $(x, y) \to (u, v)$, we apply

$$f_{U,V}(u, v) = f_{X,Y} \circ H^{-1}(u, v) \cdot |\det DH^{-1}(u, v)|,$$

where $(x, y) = H^{-1}(u, v) = (u - v, v)$ and det $DH^{-1}(u, v) = 1$.

Thus,

$$f_{U,V}(u, v) = f_{XY}(u - v, v).$$

Hence, the marginal density is

$$f_U(u) = \int_{-\infty}^{\infty} f_{XY}(u - v, v) dv.$$

Exercise 3.5

We know that

$$f_X(x) = \begin{cases} \frac{1}{3} e^{-\frac{x}{3}}, & x > 0 \\ 0, & x \leq 0 \end{cases}, \quad f_Y(y) = \begin{cases} e^{-y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

Since $U = X + Y$, then

$$\begin{aligned} f_U(u) &= \int_0^u \frac{1}{3} e^{-\frac{x}{3}} e^{-(u-x)} \mathrm{d}x \\ &= \frac{1}{3} e^{-u} \int_0^u e^{\frac{2}{3}x} \mathrm{d}x \\ &= \frac{1}{2} e^{-u} (e^{\frac{2}{3}u} - 1) \\ &= (e^{-u/3} - e^{-u})/2. \end{aligned}$$

Exercise 3.6

We assume that X_1 and X_2 are independent. Otherwise we cannot determine their covariance. We know that

$$
\begin{aligned}
f_1(x) &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \\
m_1(t) &= e^{\mu_1 t + \sigma_1^2 t^2/2} \\
f_2(x) &= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \\
m_2(t) &= e^{\mu_2 t + \sigma_2^2 t^2/2}
\end{aligned}
$$

First we prove that if $Z = \alpha X$, then $m_Z(t) = m_X(\alpha t)$.

$$
\begin{aligned}
m_Z(t) &= E[e^{tZ}] = E[e^{\alpha t X}] \\
&= E[e^{(\alpha t)X}] \\
&= m_X(\alpha t).
\end{aligned}
$$

Also, we know that if $Z = X + Y$, then $m_Z(t) = m_X(t)m_Y(t)$.

Since $Y = \lambda_1 X_1 + \lambda_2 X_2$, then

$$
\begin{aligned}
m_Y(t) =& m_{X_1}(\lambda_1 t) m_{X_2}(\lambda_2 t) \\
&= e^{\mu_1\lambda_1 t + \sigma_1^2(\lambda_1 t)^2/2 + \mu_2\lambda_2 t + \sigma_2^2(\lambda_2 t)^2/2} \\
&= e^{(\mu_1\lambda_1 + \mu_2\lambda_2)t + (\sigma_1^2\lambda_1^2 + \sigma_2^2\lambda_2^2)t^2/2}.
\end{aligned}
$$

Hence, Y follows normal distribution, where

$$
\begin{aligned}
E[Y] &= \lambda_1\mu_2 + \lambda_2\mu_2, \\
\mathrm{Var}Y &= \sigma_1^2\lambda_1^2 + \sigma_2^2\lambda_2^2.
\end{aligned}
$$

Exercise 3.7

1)

$$
\begin{aligned}
f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, x_2)\mathrm{d}x_2 \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\varrho^2)}[(\frac{x_1-\mu_1}{\sigma_1})^2 - 2\varrho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2}) + (\frac{x_2-\mu_2}{\sigma_2})^2]}\mathrm{d}x_2
\end{aligned}
$$

Since

$$
-2\varrho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2}) + (\frac{x_2-\mu_2}{\sigma_2})^2 = (\frac{x_2-\mu_2}{\sigma_2} - \varrho(\frac{x_1-\mu_1}{\sigma_1}))^2 - \varrho^2(\frac{x_1-\mu_1}{\sigma_1})^2,
$$

then

$$f_{X_1}(x_1) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\varrho^2)}(\frac{x_2-\mu_2}{\sigma_2}-\varrho(\frac{x_1-\mu_1}{\sigma_1}))^2} \mathrm{d}x_2.$$

We denote that

$$t = -\frac{1}{\sqrt{1-\varrho^2}}(\frac{x_2-\mu_2}{\sigma_2} - \varrho(\frac{x_1-\mu_1}{\sigma_1})).$$

Hence,

$$\begin{aligned} f_{X_1}(x_1) &= \frac{1}{2\pi\sigma_1} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}}. \end{aligned}$$

Hence $x_1 \sim N(\mu_1, \sigma_1^2)$.

2) We know that the coefficient of correlation of X_1 and X_2 is

$$\begin{aligned} \varrho_{X_1X_2} &= \frac{\mathrm{Cov}(X_1, X_2)}{\sigma_1\sigma_2} = \frac{E[(X_1-\mu_1)(X_2-\mu_2)]}{\sigma_1\sigma_2} \\ &= \frac{\iint_{\mathbb{R}^2}(x_1-\mu_1)(x_1-\mu_2) f_{X_1X_2}(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2}{\sigma_1\sigma_2} \end{aligned}$$

We denote $u = \frac{x_1-\mu_1}{\sigma_1}$ and $v = \frac{x_1-\mu_2}{\sigma_2}$, then

$$\begin{aligned} &\iint_{\mathbb{R}^2}(x_1-\mu_1)(x_2-\mu_2) f_{X_1X_2}(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 \\ =& \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\varrho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{1}{2(1-\varrho^2)}(u^2 - 2\varrho uv + v^2)\right] \mathrm{d}u\mathrm{d}v \\ =& \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\varrho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left[-\frac{u^2}{2} - \frac{(v-\varrho u)^2}{2(1-\varrho^2)}\right] \mathrm{d}u\mathrm{d}v \\ =& \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\varrho^2}} \int_{-\infty}^{\infty} ue^{-\frac{u^2}{2}} \left[\int_{-\infty}^{\infty} ve^{-\frac{(v-\varrho u)^2}{2(1-\varrho^2)}} \mathrm{d}v\right] \mathrm{d}u \\ =& \sigma_1\sigma_2 \int_{-\infty}^{\infty} \frac{u}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \left[\int_{-\infty}^{\infty} \frac{v}{\sqrt{2\pi}\sqrt{1-\varrho^2}} e^{-\frac{(v-\varrho u)^2}{2(1-\varrho^2)}} \mathrm{d}v\right] \mathrm{d}u \\ =& \sigma_1\sigma_2 \int_{-\infty}^{\infty} \frac{u}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \cdot \varrho u \ \mathrm{d}u \\ =& \varrho\sigma_1\sigma_2 \int_{-\infty}^{\infty} \frac{u^2}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \mathrm{d}u \\ =& \varrho\sigma_1\sigma_2. \end{aligned}$$

Hence,

$$\varrho_{X_1X_2} = \frac{\varrho\sigma_1\sigma_2}{\sigma_1\sigma_2} = \varrho.$$

3) Sufficiency:

if $\varrho = 0$, then

$$f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}[(\frac{x_1-\mu_1}{\sigma_1})^2+(\frac{x_2-\mu_2}{\sigma_2})^2]}.$$

Also,

$$\begin{aligned} f_{X_1}(x_1)f_{X_2}(x_2) &= \frac{1}{2\pi\sigma_1\sigma_2}e^{[-(\frac{x_1-\mu_1}{2\sigma_1}^2+\frac{x_2-\mu_2}{2\sigma_2})^2]} \\ &= f_{X_1X_2}(x_1,x_2), \end{aligned}$$

which means X_1 and X_2 are independent.

Necessity:

if X_1 and X_2 are independent, then

$$f_{X_1X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2).$$

Specially, when $x_1 = \mu_1$ and $x_2 = \mu_2$,

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{2\pi}\sigma_2}.$$

Hence $\varrho = 0$.

It's not true for a bivariate random variable with an arbitrary distribution.

For example,

X,Y	0	$\frac{1}{4}$	1
-1	0	0	$\frac{1}{5}$
$\frac{1}{2}$	0	$\frac{1}{5}$	0
0	$\frac{1}{5}$	0	0
$\frac{1}{2}$	0	$\frac{1}{5}$	0
1	0	0	$\frac{1}{5}$

Here we know that $Y = X^2$. However,

$$
\begin{aligned}
&E[XY] = -\frac{1}{5} + \frac{1}{8}\frac{1}{5} + 0\frac{1}{5} + \frac{1}{5} - \frac{1}{8}\frac{1}{5} = 0,\\
&E[X]E[Y] = 0, \quad E[Y] = 0.\\
&\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = 0.
\end{aligned}
$$

Here $\varrho_{XY} = 0$, but X and Y are not independent.

4) First, we derive the conditional density

$$
f_{X_2|x_1} = \frac{f_{X_1X_2}(x_1,x_2)}{f_{X_1}(x_1)}.
$$

Then, we denote $u = \frac{x_1-\mu_1}{\sigma_1}$ and $v = \frac{x_1-\mu_2}{\sigma_2}$.

$$
\begin{aligned}
\mu_{X_2|x_1} &= E[X_2|_{x_1}] = \int_{-\infty}^{\infty} x_2 f_{X_2|x_1}(x_2)\mathrm{d}x_2\\
&= \sigma_2 \int_{-\infty}^{\infty} \frac{v\sigma_2+\mu_2}{\sqrt{2\pi}\sigma_2\sqrt{1-\varrho^2}} e^{-\frac{1}{2(1-\varrho^2)(u^2-2\varrho uv+v^2)}} e^{-\frac{u^2}{2}}\mathrm{d}v\\
&= \int_{-\infty}^{\infty} \frac{v\sigma_2+\mu_2}{\sqrt{2\pi}\sqrt{1-\varrho^2}} e^{-\frac{u^2}{2}-\frac{(v-\varrho u)^2}{2(1-\varrho^2)}} e^{\frac{u^2}{2}}\mathrm{d}u\\
&= \int_{-\infty}^{\infty} \frac{v\sigma_2+\mu_2}{\sqrt{2\pi}\sqrt{1-\varrho^2}} e^{-\frac{(v-\varrho u)^2}{2(1-\varrho^2)}}\mathrm{d}u\\
&= \sigma_2 \int_{-\infty}^{\infty} \frac{v}{\sqrt{2\pi}\sqrt{1-\varrho^2}} e^{-\frac{(v-\varrho u)^2}{2(1-\varrho^2)}}\mathrm{d}u + \mu_2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1-\varrho^2}} e^{-\frac{(v-\varrho u)^2}{2(1-\varrho^2)}}\mathrm{d}u\\
&= \sigma_2\varrho u + \mu_2\\
&= \mu_2 + \varrho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1).
\end{aligned}
$$

Exercise 3.8

1) unit:10.

Stem	Leaves
532	9
534	2
535	47
536	6
537	5678
538	12345778888
539	016999
540	11166677889
541	123666688
542	0011222357899
543	01111556
544	00012455678
545	233447899
546	23569
547	357
548	11257

We can approximately conclude that it follows a shape of normal distribution.

2) We know that n=100, so that the number of categories is 8.

The data range is 5487-5329=158. The category length is $\lceil 158/8 \rceil = 20$.

The lower bound is $5329 - 0.5 = 5328.5$.

```
Histogram[Data, Min[cat], Max[cat], 20, PlotRange -> 0, Automatic,
Frame -> True, True, False, False,
FrameTicks -> Automatic, None, cat, None,
FrameLabel -> "Shear strength of spot welds", "Number of welds"]
```

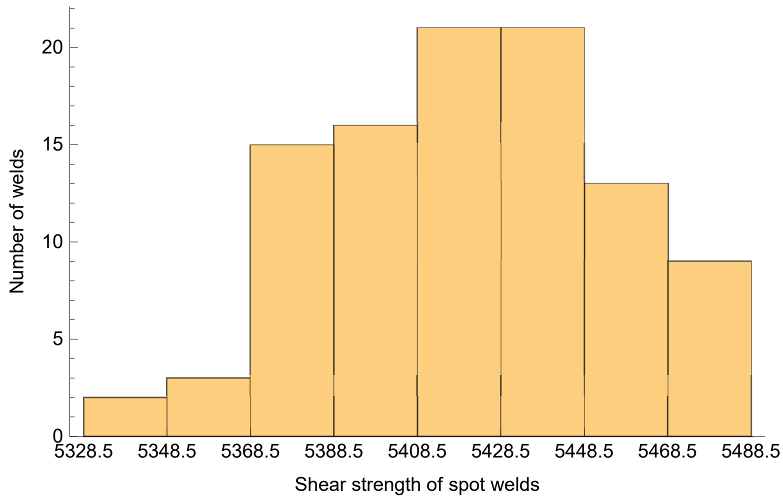

Figure 1: Shear strength distribution of 100 sample welds

The histogram obtains a shape similar to normal distribution, conveying the same information as the stem-leaf display.

3)

$$q_1 = 5399 \quad q_2 = 5421.2 \quad q_3 = 5445.5.$$
$$iqr = 46.5, \quad f_1 = 5329.25, \quad f_3 = 5515.25.$$
$$a_1 = 5342, \quad a_3 = 5487.$$

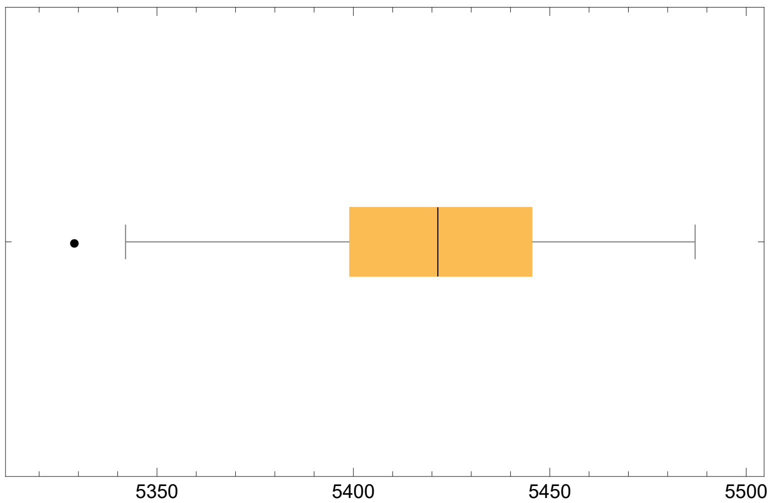

Figure 2: Boxplot of shear strength of 100 sample welds

Boxplot demonstrates the distribution more simply and clearly, stating the key points without specific data values. Stem and leaf diagram is more complicated but keeps all the details.

Exercise 3.9

1) unit:1.

Stem	Leaves
10	1
22	588
23	02234
23	558
24	023

2) Sample mean: 22.16.

   Sample median: 23.35.

   Sample standard deviation: 3.47.

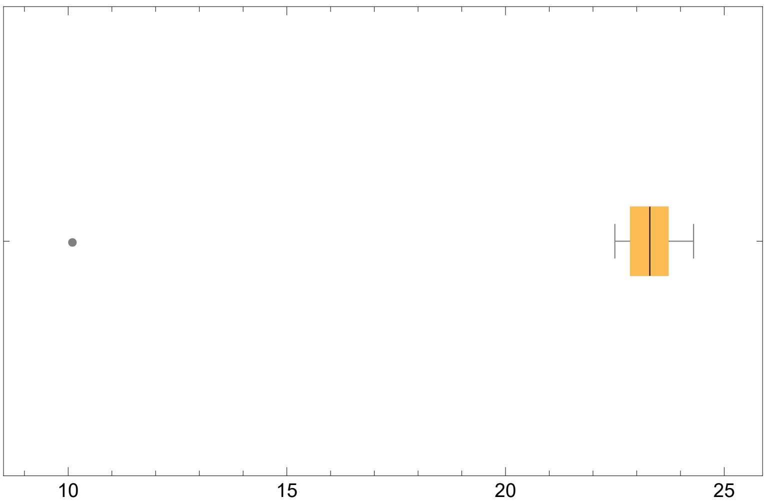

Figure 3: Boxplot of temperature samples

3)

4) Sample mean: 23.39.

   Sample median: 23.35.

Sample standard deviation: 0.54.

The sample mean is more accurate after we drop the outlier. The sample standard deviation is much less than before because the degree of dispersion is less.

Exercise 3.10

$$
\begin{aligned}
L(\gamma) &= \prod_{i=1}^{n} f(X_i) = \prod_{i=1}^{n} (\gamma+1)x_i^{\gamma} \\
&= (\gamma+1)^n (\prod_{i=1}^{n} x_i)^{\gamma}.
\end{aligned}
$$

We take the logarithm,

$$
\begin{aligned}
\ln L(\gamma) &= n\ln(\gamma+1) + \gamma \ln(\prod_{i=1}^{n} x_i) \\
&= n\ln(\gamma+1) + \gamma \sum_{i=1}^{n} \ln(x_i).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\gamma} L(\gamma) &= \frac{n}{\gamma+1} + \sum_{i=1}^{n} \ln(x_i) = 0. \\
\gamma &= -\frac{1}{n\sum_{i=1}^{n} \ln(x_i)} - 1.
\end{aligned}
$$

Exercise 3.11

1) We know that the density function is

$$
f(x) = \binom{4}{x} p^x (1-p)^{4-x}, \quad x = 0, 1, 2, 3, 4.
$$

Thus,

$$
L(p) = \left[\binom{4}{0} p^0 (1-p)^4\right]^8 \left[\binom{4}{1} p^1 (1-p)^3\right]^6 \left[\binom{4}{2} p^2 (1-p)^2\right] = 24p^8(1-p)^{52}.
$$

We take the derivative,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p} L(p) &= 24(8p^7(1-p)^{52} - p^8 \cdot 52(1-p)^5 1) = 0, \\
p &= \frac{2}{15}.
\end{aligned}
$$

2) The expectation of failed brake shoes is

$$E[X] = np = 4 \cdot \frac{2}{15} = \frac{8}{15}.$$

Hence the failure rate is

$$p = \frac{E[X]}{n} = \frac{2}{15} \approx 0.133 > 0.1.$$

Hence, I will have some doubts concerning the use of this new material.