Stats Sample 2.

1. a) i) unimodal, fairly ~~symmetric~~ symmetric, bell-shaped.

ii) A) $\bar{x} \pm t_{57;0.995} \frac{s}{\sqrt{n}}$

$= 70.7 \times 10^3 \pm 2.678 \times \frac{1.78 \times 10^3}{\sqrt{58}}$

$= 70.7 \times 10^3 \pm 0.626 \times 10^3$

$= [70.074 \times 10^3 \quad 71.326 \times 10^3]$

B) Data is randomly selected. (independent) (Cannot be checked)
Data come from a normal distributed population.
(drawing a quantile plot can check this)

iii) A) $\hat{p} = \frac{21}{58} \approx 0.362$

B) $p = \hat{p} \pm z_{0.975} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= \frac{21}{58} \pm 1.96 \sqrt{\frac{\frac{21}{58}(1-\frac{21}{58})}{58}}$

$= \frac{21}{58} \pm 0.2574$

$= [0.1047 \quad 0.6195]$

C) Normality assumption.
Random assumption.
n ~~is~~ should be big enough. i.e. $n\pi(1-\pi) > 5$.
$58 \times 0.1047(1-0.1047) = 5.437 > 5$
∴ assumption is ~~reson~~ reasonable.

b) i) P(have file) $= \frac{10}{15} = \frac{2}{3}$
Let X be no. of CDs that have files. $X \sim Bin(12, \frac{2}{3})$
$P(X=9) = 0.212$

ii) $0.212 \times \frac{1}{3} = 0.0707$

Q2.

a). Random Assumption: Sample is collected randomly and data is ~~inde~~ independent

Normality Assumption: Sample is collected from a normally distributed population.

Constant ~~standard deviation~~ variance assumption.

i.e. σ is the same for all ~~so data~~. groups.

b).

Source	df	SS	MS	F
Treatment	2	0.0608	0.0304	0.99
Error	12	0.3701	0.03084	
Total	14	0.4309		

c) Let μ_1, μ_2, μ_3 denote the true mean degree of soiling for those 3 different mixtures.

$H_0: \mu_1 = \mu_2 = \mu_3$ against H_a: Not all means equal

observation Value $f_0 = 0.99$, following distribution $F_{2;12}$.

Reject H_0 if $f_0 > f_{2;12;0.95}$

$f_{2;12;0.95} = 3.89 > 0.99$. (table)

$\therefore$ Do not reject H_0.

OR. Reject H_0 if P-value < 0.05.

P-value $= P(X > 0.99) > 0.05$ (table).

$\therefore$ Do not Reject H_0.

Therefore, There is not enough evidence to suggest that the true mean degree of soiling differs amoung different mixtures. That is, the three ~~m~~ different mixtures can have the same ~~def deg~~ mean degree of soiling.

d) $CI = \bar{X}_2 - \bar{X}_3 \pm t_{12;0.975}\sqrt{MS_{Er}\left(\frac{1}{n_2}+\frac{1}{n_3}\right)}$

$= 0.794 - 0.938 \pm 2.179\sqrt{0.03084\left(\frac{1}{5}+\frac{1}{5}\right)}$

$= -0.144 \pm 0.242$.

$= [-0.386, 0.098]$

No. because 0 is contained in the CI.

Q2.

e). ~~Although analysing~~ $H_0: \mu_1 = \mu_2$, $\mu_2 = \mu_3$, $\mu_3 = \mu_1$ against H_a: Not all equal

$0.14 > \frac{0.05}{3}$ → to get overal $\alpha = 0.05$, must use $\frac{\alpha}{K}$ ← $\binom{k}{2}$

$0.86 > \frac{0.05}{3}$

CI in d contains 0.

∴ Do not reject $H_0: \mu_1 = \mu_2$, $\mu_2 = \mu_3$, $\mu_1 = \mu_3$.

~~Although~~ the three pairwise comparisons give the same result (i.e. Do not reject), ~~This does not always work.~~

Q3. a) i) A) Random assumption (independence): residuals must be random and independent.

Normal assumption: residuals must come from normal distributed population.

equal variance ~~Constant σ~~ assumption: common variance for different sample.

B) ~~Random~~ Dots in right graph shows no pattern for the residual. That means the equal variance assumption stands.

Dots in left graph does not seem to be random ~~distribu~~

This may mean that the true population is not normally distributed, so the independence assumption does not stand.

$\boxed{\text{ii) } r^2 = 0.6325}$

iii) $H_0: \beta_1 = 0$. against $H_a: \beta_1 \neq 0$.

observed value $t_0 = -3.711$, following distribution t_8

Reject H_0 if ~~$t_{8;0.975}$~~ p-value < 0.05

p-value $= 2 \times P(T > 3.711)$ $T \sim t_8$.

$> 2 \times 0.005$

$= 0.01 < 0.05$.

∴ Reject H_0

OR. ~~$t_{8;0.975} = 2.306$~~

~~$-3.711 < -2.306$~~

~~∴ Reject H_0.~~

Therefore, applied stress has ~~there is~~ significant impact to the time to failure.

iv) $\beta_1 = \hat{b}_1 \pm t_{8;0.975} \cdot \text{SE coeff}$.

$= \cancel{2.2428}\ -0.9009 \pm 2.306 \times 0.2428$

$= -0.9009 \pm 0.5599 = [-1.4608, -0.341]$

3.

b) i) $E[Y_1 \mid X = 25] = 65 - 1.2 \times 25 = 35.$

$E[Y_2 \mid X = 24] = 65 - 1.2 \times 24 = 36.2$

$Y_1 \sim N(35, \sqrt{8}) \qquad Y_2 \sim N(36.2, \sqrt{8})$

$Y_1 - Y_2 = N(-1.2, 4)$

ii) A) $P(Y_1 > Y_2) = P(Y_1 - Y_2 > 0)$

$= 1 - P(Y_1 - Y_2 < 0)$

$= 1 - P\left(Z < \frac{0 + 1.2}{4}\right)$

$= 1 - \cancel{0.5319}\ 0.6179$

$= \cancel{0.4681}\ 0.3821$

B) The possibility of $Y_1 > Y_2$ is less than $\frac{1}{2}$.
This means Y_1 is less likely to be higher than Y_2.
Therefore, as x decreases, Y_1 is more likely to increase.