Answers to November 2008 exam

1. (1-1) D; [$0.6{\times}0.02 + 0.3{\times}0.03 + 0.1{\times}0.04 = 0.025$]

   (1-2) B; [$\frac{0.004}{0.025} = 0.16$]

   (1-3) E; [$p_{22}^{(2)} = 0.33$]

   (1-4) E; [$\mathrm{var}(Z) = 4^2 + 3^2$]

   (1-5) C; [the Normal distribution has shorter tails than exponential; it must therefore have an increasing hazard function.]

   (1-6) A; [$(Z > 1) \cap (Z > 2) = (Z > 2)$, since $(Z > 2) \subset (Z > 1)$]

   (1-7) C; [by definition]

   (1-8) B; [$z = \frac{15.5-15}{0.204} = 2.45 \;\Rightarrow\; \mathrm{p} = 0.014$]

   (1-9) D; [$\mathrm{RLL} \geqslant -2 \;\Rightarrow\; -3 < \theta < 15$]

   (1-10) A; [$df = (3, 3, 3, 6, 15)$; $F = \frac{40}{10} = 4$, cf. $\mathrm{F}_{3,6}$]

2. (a) i. $\Pr(F \cup G) = \Pr(F) + \Pr(G) \;\Rightarrow\; 0.8 = 0.2 + g \;\Rightarrow\; g = 0.6$;

   ii. $\Pr(F \cup G) = \Pr(F) + \Pr(G) - \Pr(F \cap G) \;\Rightarrow\; 0.8 = 0.2 + g - 0.2g \;\Rightarrow\; g = 0.75$;

   iii. probability table:

	G	G'	
F	$g{-}0.6$	$0.8{-}g$	0.2
F'	0.6	0.2	0.8
	g	$1{-}g$	1

   $\Pr(F \mid G) = \frac{g-0.6}{g}$; and $g \leqslant 0.8$, so $\Pr(F \mid G) \leqslant \frac{0.2}{0.8} = 0.25$.

   (b) $X \overset{\mathrm{d}}{=} \mathrm{Bi}(n{=}20, p{=}0.1)$, so $\Pr(X \leqslant 2) = 0.1216 + 0.2702 + 0.2852 = 0.677$ (Tables).

   (c) i. $P = \begin{bmatrix} 0.2 & 0.8 & 0 & 0 \\ 0.1 & 0.1 & 0.8 & 0 \\ 0.1 & 0 & 0.1 & 0.8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$;

   ii. five-step transition probabilities are given by P^5; $p_{14}^{(5)}$, the (1,4)-element of P^5, gives the probability of going from state 1 to state 4 in five steps, i.e. the probability of completion in 5 hours.

3. (a) $\mathrm{E}(X) = 0{\times}0.2 + 1{\times}0.6 + 2{\times}0.2 = 1$;
   $\mathrm{E}(X^2) = 0{\times}0.2 + 1{\times}0.6 + 4{\times}0.2 = 1.4 \;\Rightarrow\; \mathrm{var}(X) = 0.4$.
   $\therefore\; \mathrm{E}(Y) = 160{\times}1 = 160$, $\mathrm{var}(Y) = 160{\times}0.4 = 64$.
   $\Pr(Y \leqslant 150) \approx \Pr(Y^* < 150.5) = \Pr(Y_s^* < -1.188) = 0.118$.

   (b) $F(t) = 1 - e^{-H(t)}$, where $H(t)$ denotes the cumulative hazard function.
   $H(4) = \int_0^{0.4} 0.03\sqrt{t}dt = [\, 0.03\frac{4^{3/2}}{3/2} \,] = 0.16$;
   $\therefore\; \Pr(T > 4) = e^{-H(4)} = e^{-0.16} = 0.852$.

   (c) approx formula with $\psi(x) = \sqrt{x}$, so $\psi'(x) = \frac{1}{2\sqrt{x}}$, $\psi''(x) = -\frac{1}{4x\sqrt{x}}$; $\mu = \theta, \sigma^2 = c\theta^2$;
   $\mathrm{E}(\sqrt{Z}) \approx \sqrt{\theta} + \frac{1}{2}(-\frac{1}{4\theta\sqrt{\theta}})c\theta^2 = \sqrt{\theta} - \frac{1}{8}c\sqrt{\theta} = \sqrt{\theta}(1 - \frac{c}{8})$;
   $\mathrm{var}(\sqrt{Z}) \approx (\frac{1}{2\sqrt{\theta}})^2 c\theta^2 = \frac{c\theta}{4}$.
   The approximations are best when $\mathrm{var}(Z)$ is small, i.e. if c is small.

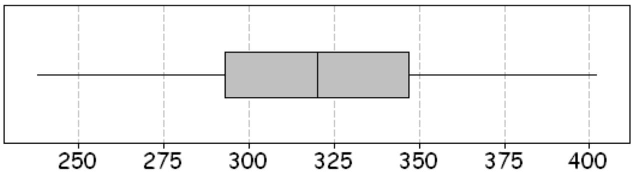
   (d) i. $\Pr(X > 50) = \Pr(X_s > 1) = 0.159$;

   ii. $X - T \overset{\mathrm{d}}{=} \mathrm{N}(-4, 5^2)$, since $46{-}50 = -4$ and $3^2{+}4^2 = 5^2$.
   $\Pr(X > T) = \Pr(X{-}T > 0) = \Pr((X{-}T)_s > 0.8) = 0.212$.

4. (a) $\bar{x} = 27.14$, $s = 9.85$;

   (b) $\min \sim c_{0.02} = 320 - 2.0537{\times}40 = 237.9$;
   $\mathrm{Q1} \sim c_{0.25} = 320 - 0.6745{\times}40 = 293.0$;
   $\mathrm{med} \sim c_{0.5} = 320$;
   $\mathrm{Q1} \sim c_{0.75} = 320 + 0.6745{\times}40 = 347.0$;
   $\max \sim c_{0.98} = 320 + 2.0537{\times}40 = 402.1$;

five-number summary: (238, 293, 320, 347, 402).



(c) "x-coordinate" = $\Phi^{-1}(\frac{15}{20}) = 0.6745$; "$y$-coordinate" = $x_{(15)} = 65$.
y-intercept $\Rightarrow$ $\hat{\mu} = 50$; slope $\Rightarrow$ $\hat{\sigma} = 20$ $(= \frac{70-50}{1-0})$

5. $\bar{X} \overset{\text{d}}{=} \text{N}(\mu, \frac{5^2}{12})$

i. 95% probability interval for $\bar{X}$: $40 \pm 1.96 \times \frac{5}{\sqrt{12}} = (40 \pm 2.83) = (37.17, 42.83)$.

ii. size $= \Pr(|\bar{X} - 40| > 3)$, where $\bar{X} \overset{\text{d}}{=} \text{N}(40, \frac{5^2}{12})$;
$\therefore$ size $= 2\Pr(\bar{X}_s > \frac{3\sqrt{12}}{5}) = 2\Pr(\bar{X}_s > 2.078) = 0.038$.

iii. $\text{p} = 2\Pr(\bar{X} > 44.0)$, where $\bar{X} \overset{\text{d}}{=} \text{N}(40, \frac{5^2}{12})$;
$\therefore$ $\text{p} = 2\Pr(\bar{X}_s > \frac{4\sqrt{12}}{5}) = 2\Pr(\bar{X}_s > 2.771) = 0.006$.

iv. power $= \Pr(|\bar{X}' - 40| > 3)$, where $\bar{X}' \overset{\text{d}}{=} \text{N}(45, \frac{5^2}{12})$;
$\therefore$ power $= 1 - \Pr(37 < \bar{X}' < 43) = 1 - \Pr(-5.543 < \bar{X}'_s < -1.386) = 0.917$.

v. 95% confidence interval for μ: $44.0 \pm 1.96 \times \frac{5}{\sqrt{12}} = (44.0 \pm 2.83) = (41.17, 46.83)$.

vi. 95% prediction interval for X: $44.0 \pm 1.96 \times 5\sqrt{\frac{13}{12}} = (44.0 \pm 10.2) = (33.8, 54.2)$.

6. (a) $\hat{\lambda} = \bar{x} = 5.7$; $\sigma^2 = \text{var}(X) = \lambda^2$, so $\text{var}(\bar{X}) = \frac{\lambda^2}{10}$;
$\text{se}(\hat{\lambda}) = \frac{\hat{\lambda}}{\sqrt{10}} = \frac{5.7}{\sqrt{10}} = 1.8$.

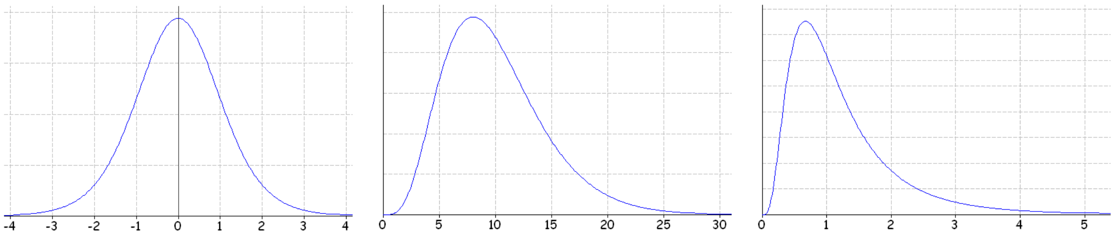
(b) $\frac{\partial \ln L}{\partial \theta} = -20 + \frac{100}{\theta} = 0 \Rightarrow \hat{\theta} = 5$;
$-\frac{\partial^2 \ln L}{\partial \theta^2} = \frac{100}{\theta^2} \Rightarrow \text{se} = 1/\sqrt{\frac{100}{25}} = 0.5$.

(c) i. tables: $\text{E}(Z_{(12)}) = 1.629 \Rightarrow \text{E}(R) = 2 \times 1.629\sigma$;
$\hat{\sigma} = \frac{\bar{R}}{3.258} = \frac{7.32}{3.258} = 2.25$.

ii. control limits for $\bar{X}$: $36.2 \pm 3 \times \frac{2.25}{\sqrt{12}} = (36.29 \pm 1.95) = (34.34, 38.24)$.

iii. tables: $\text{E}(Z_{(11)}) = 1.116 \Rightarrow \text{E}(Q) = 2 \times 1.116\sigma$; thus $c = 2.232$.
The estimator $\hat{\sigma}^* = \frac{\bar{Q}}{2.232}$ is more stable than $\hat{\sigma} = \frac{\bar{R}}{3.258}$. It is less susceptible to outliers, which is both good and bad. It may be important to pick up outliers; but it is important to have a good estimate of the underlying variation.

7. (a) t_{10} pdf χ^2_{10} pdf $F_{10,10}$ pdf



(b) i. tables: $\Pr(0.226 < \text{F}_{8,8} < 4.433) = 0.95$,
95% CI for $\frac{\sigma_1^2}{\sigma_2^2}$: $\left(0.226 \times \frac{7.41^2}{5.67^2} < \frac{\sigma_1^2}{\sigma_2^2} < 4.433 \times \frac{7.41^2}{5.67^2}\right) = (0.385 < \frac{\sigma_1^2}{\sigma_2^2} < 7.571)$;
$\therefore$ 95% confidence interval for $\frac{\sigma_1}{\sigma_2}$: $(0.62 < \frac{\sigma_1}{\sigma_2} < 2.75)$.
Since $1 \in$ CI, there is no significant evidence against $\sigma_1 = \sigma_2$ in these data.

ii. $s^2 = \frac{1}{2}(s_1^2 + s_2^2)$, since the sample sizes are equal;
hence $s = \sqrt{\frac{1}{2}(7.41^2 + 5.47^2)} = \sqrt{43.53} = 6.60$.

iii. $\frac{16S^2}{\sigma^2} \overset{\mathrm{d}}{=} \chi^2_{16} \;\Rightarrow$

95% confidence interval for σ: $\left(\sqrt{\frac{16\times 43.53}{28.85}},\ \sqrt{\frac{16\times 43.53}{5.908}}\right) = (4.91, 10.04)$

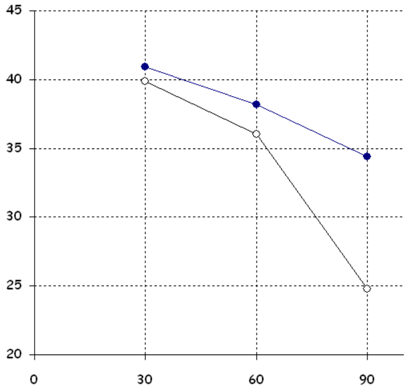
iv. 95% confidence interval for $\mu_1 - \mu_2$: $7.4 \pm 2.120\sqrt{43.53(\frac{1}{9}+\frac{1}{9})}$
$= (7.4 \pm 6.59) = (0.81, 13.99)$.

8. (a) i. error MS = $\mathrm{av}(S^2) = \frac{1}{3}(20+24+31) = 25$

treatment	2	200	100	4
error	15	375	(25)	
total	17	(575)		

ii. $H_0 \;\Rightarrow\; F \overset{\mathrm{d}}{=} \mathrm{F}_{2,15}$, $c = c_{0.95}(\mathrm{F}_{2,15}) = 3.7$;
$F = 4$, and since $F > c$, we reject H_0. There is evidence of significant treatment effects. (It appears that both A and B are significantly greater than the control.)

iii. $c_{0.95}(\mathrm{F}_{1,6}) = 5.99$, $c_{0.95}(\mathrm{F}_{2,6}) = 5.14$; so all effects are significant: type, humidity and interaction are all significant. The cell mean plot is shown below:

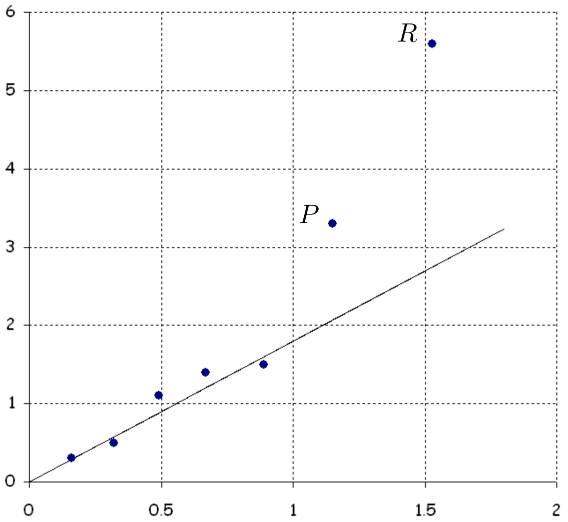


Humidity reduces the required force; A is better than B (greater required force); humidity has a greater effect on B (it reduces the required force by more).

9. (a) The "full" model is fitted, which has as many parameters as degrees of freedom and therefore gives a "perfect" fit, with no degrees of freedom left over for error, and zero error, so that error SS = 0.

(b) RMS = 3.3, 1.5, 5.6, 1.4, 0.5, 1.1, 0.3;

ordered RMS =	0.3,	0.5,	1.1,	1.4,	1.5,	3.3,	5.6;
half-normal quantile =	0.16,	0.32,	0.49,	0.67,	0.89,	1.15,	1.53.



It appears that P and R are significant, and the others are not.

(c) Assume that the effects other than P and R are zero, so that their sums of squares can be attributed to error.

P	1	11.1	$F = 9.4^*$
R	1	31.2	$F = 26.4^*$
e	5	5.9	$s^2 = 1.18$
total	7	48.2	

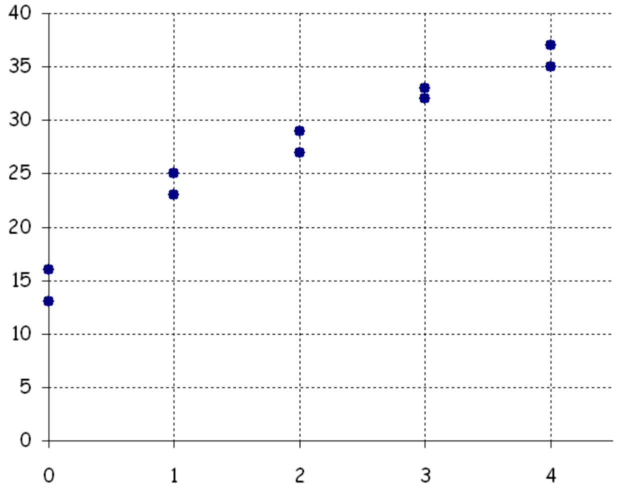
Since $c_{0.95}(\mathrm{F}_{1,5}) = 6.61$, these results indicate that P and R are both significant.

(d) $\hat{R} = -4.0$ (= av(R1) – av(R0) = 13.9 – 17.9);
95% CI for R: $-4.0 \pm 2.571\sqrt{1.18(\frac{1}{4} + \frac{1}{4})} = (-4.0 \pm 1.97) = (-5.97, -2.03)$.

(e) P has a significant (positive) effect; R has a significant (negative) effect; there is no evidence that Q has an effect, and nor is there evidence of any interactions.

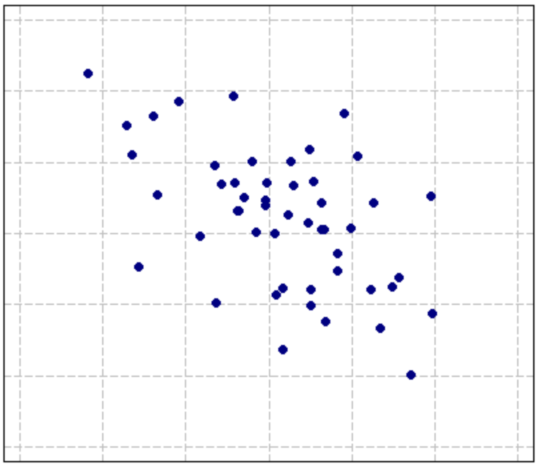
10. (a) i. $\hat{\eta}_2 = \bar{y} = 27.0$, since $\hat{\eta}_x = \bar{y} + \hat{\beta}(x - \bar{x})$;
95% confidence interval: $27.0 \pm 2.306\sqrt{\frac{4.44}{10}} = (27.0 \pm 1.54) = (25.46, 28.54)$.

ii. $\hat{\mu}_2 = \bar{y}_2 = 28.0$;
95% confidence interval: $28.0 \pm 2.571\sqrt{\frac{2.20}{2}} = (28.0 \pm 2.70) = (25.30, 30.70)$.



The straight-line regression assumption looks dubious: the observations at $x = 0$ don't seem to fit.

(b) i. scatter plot with $\rho \approx -0.5$:



ii. Table 20: $(-0.68 < \rho < -0.26)$.