THE ROYAL STATISTICAL SOCIETY

2005 EXAMINATIONS – SOLUTIONS

HIGHER CERTIFICATE

PAPER III
STATISTICAL APPLICATIONS AND PRACTICE

The Society provides these solutions to assist candidates preparing for the examinations in future years and for the information of any other persons using the examinations.

The solutions should NOT be seen as "model answers". Rather, they have been written out in considerable detail and are intended as learning aids.

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

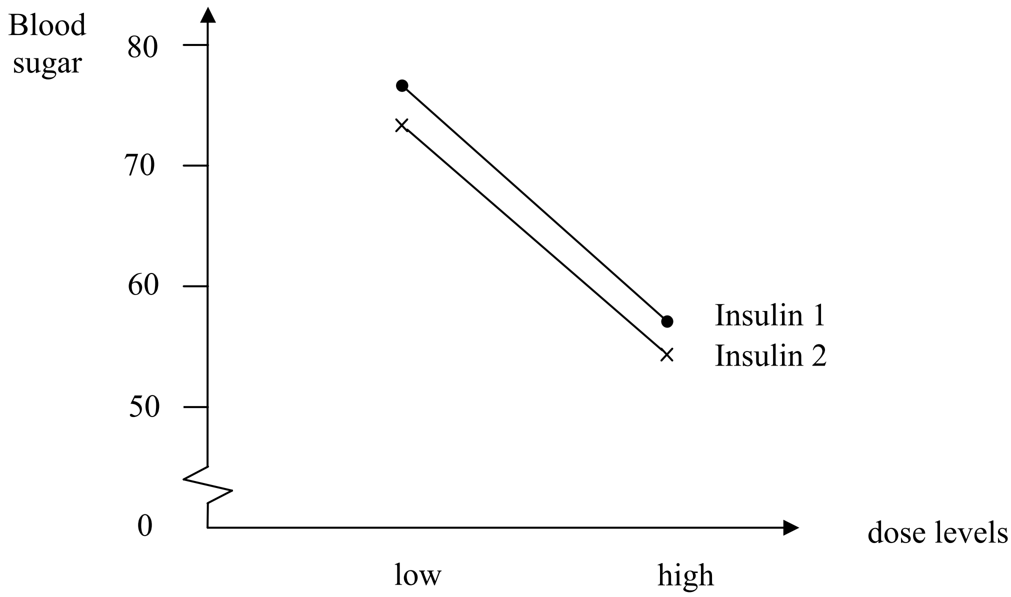
While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of these solutions.

Note. In accordance with the convention used in the Society's examination papers, the notation log denotes logarithm to base *e*. Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

<u>Higher Certificate, Paper III, 2005. Question 1</u>

(i) Means are: I low 76.25; I high 57.50; II low 73.75; II high 54.25.



As the two lines are virtually parallel, there is no evidence of any interaction between insulin type and dose level.

(ii) Totals for insulins are I: 535, II: 512. Grand total = 1047

Hence SS for insulin $= \dfrac{535^2}{8} + \dfrac{512^2}{8} - \dfrac{1047^2}{16} = 33.0625$.

Level totals are 600, 447. So SS for levels $= \dfrac{600^2}{8} + \dfrac{447^2}{8} - \dfrac{1047^2}{16} = 1463.0625$.

So analysis of variance table is

SOURCE	DF	SS	MS	*F* value
Rabbits	3	297.19	99.06	1.33 compare $F_{3,9}$
Insulin	1	33.06	33.06	0.444 compare $F_{1,9}$
Dose level	1	1463.06	1463.06	19.65 …
Insulin × Dose	1	0.57	0.57	0.008 …
Treatments	3	1496.69		
Residual	9	670.06	74.45	$= \hat{\sigma}^2$
TOTAL	15	2463.94		

The standard error of a treatment mean is $\sqrt{\dfrac{74.45}{4}} = 4.31$

Solution continued on next page

(iii) The only influential effect on blood sugar is the dose given; there is no evidence of any differences due to types of insulin and certainly not of any interaction of dose level with type of insulin. The higher dose level reduces blood sugar. Results are rather variable, as shown by the size of the standard error. Rabbits do not show any real difference in response.

(iv) Using the same four rabbits for all treatments eliminates any possible differences between animals (which did not show up in this experiment but may do in others). Treatment effects and differences will be estimated more precisely because of this. But we need to assume that using the same animals for all four treatments does not affect the responses, all of which are still independent of one another. If there were to be reactions or carry-over effects, it would be better to use 16 animals. The results would be obtained more quickly but they would very likely be more variable.

Higher Certificate, Paper III, 2005. Question 2

(i) (a) If we assume that the two samples of cars are drawn at random from populations having the same variance, and petrol consumption can be assumed Normally distributed, the means of *A* and *B* can be compared by a two-sample *t* test.

The null hypothesis is that the two population means are the same, and the alternative is that mean *A* > mean *B*. So this is a one-sided test.

$$n_A = 6, \quad n_B = 6\,;$$

$$\bar{x}_A = 8.652, \quad \bar{x}_B = 8.335;$$

$$s_A^2 = 2.2865\,, \quad s_B^2 = 1.8578.$$

To test for equality of population variances, consider $s_A^2/s_B^2 = 1.23$. This is not significant as an observation from $F_{5,5}$ so it is reasonable to take the population variances as equal.

The pooled estimate of the common variance is $s^2 = 2.0722$, with 10 d.f.

Thus the test statistic for testing $\mu_A = \mu_B$ is

$$\frac{\bar{x}_A - \bar{x}_B(-0)}{s\sqrt{\frac{1}{6}+\frac{1}{6}}} = \frac{0.317}{0.831} = 0.38\,,$$

which is referred to t_{10}. This is not significant at the 5% level (upper single-tailed 5% point is 1.812), so there is no evidence to reject the null hypothesis – it seems that the population means are the same.

Solution continued on next page

(b) Assuming the samples will be reasonably large, we would base the test on the Normal distribution and use test statistic $z = \dfrac{\bar{x}_A - \bar{x}_B}{s\sqrt{\frac{1}{n}+\frac{1}{n}}}$ where n is the size of each sample and s will be taken as $\sqrt{2.0722} = 1.439$. The null hypothesis is rejected (at the 5% level) if the value of z is >1.645. We wish to have probability 0.95 that this will happen if in fact $\mu_A - \mu_B = 0.5$. Thus we require

$$0.95 = P\left(\frac{\bar{X}_A - \bar{X}_B}{s\sqrt{\frac{2}{n}}} > 1.645 \mid \mu_A - \mu_B = 0.5\right)$$

$$= P\left(\bar{X}_A - \bar{X}_B > 1.645 s\sqrt{\tfrac{2}{n}} \mid \mu_A - \mu_B = 0.5\right).$$

The underlying distribution of $\bar{X}_A - \bar{X}_B$ is taken as $\mathrm{N}\left(\mu_A - \mu_B\, ,\, \left[s\sqrt{\tfrac{2}{n}}\right]^2\right)$, i.e. here it is $\mathrm{N}\left(0.5\, ,\, \left[s\sqrt{\tfrac{2}{n}}\right]^2\right)$. So we require

$$0.95 = P\left(\mathrm{N}\left(0.5\, ,\, \left[s\sqrt{\tfrac{2}{n}}\right]^2\right) > 1.645 s\sqrt{\tfrac{2}{n}}\right) = P\left(\mathrm{N}(0,1) > \frac{1.645 s\sqrt{\frac{2}{n}} - 0.5}{s\sqrt{\frac{2}{n}}}\right).$$

But $0.95 = P(\mathrm{N}(0, 1) > -1.645)$.

So $-1.645 s\sqrt{\tfrac{2}{n}} = 1.645 s\sqrt{\tfrac{2}{n}} - 0.5$ leading to

$$0.5 = 3.29 s\sqrt{\tfrac{2}{n}} \quad \text{i.e.} \quad \sqrt{n} = \frac{3.29 s\sqrt{2}}{0.5} \quad \text{i.e.} \quad n = \frac{(3.29)^2 \times 2.0722 \times 2}{(0.5)^2},$$

i.e. $n = 179.44$. Thus $n = 180$ is required for each group, or 360 in total.

(ii) (a) We can now eliminate any systematic differences between cars, and hope to obtain more precise results.

In each pair, the order of using with/without should be chosen at random to avoid time differences. Perhaps the same driver could be used for all, or at least limit the number and design the experiment to balance order and drivers.

(b) We use the paired-sample t test for the differences (6.55 – 6.15 etc); these are 0.40, 1.05, 0.46, 0.89, 0.08, 1.36. We assume that these differences are a sample from a Normal distribution.

We have $\bar{x}_d = 0.707$ and $s_d^2 = 0.2252$. Thus the test statistic is

$$\frac{\bar{x}_d - 0}{s_d / \sqrt{5}} = \frac{0.707}{0.194} = 3.65,$$

which is referred to t_5. This is significant at the 1% level (upper single-tailed 5% point is 3.365), so there is strong evidence to reject the null hypothesis – it seems that the population means are not the same and that the additive is beneficial.

Higher Certificate, Paper III, 2005. Question 3

(i) Define $S_{xy} = \sum xy - \frac{\sum x \sum y}{n}$, etc. Then $\hat{\beta} = \frac{S_{xy}}{S_{xx}}$ and $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$.

We have $n = 38$, so $\bar{x} = \frac{2436}{38} = 64.11$ and $\bar{y} = \frac{1670}{38} = 43.95$.

$S_{xy} = 116888 - \frac{2436 \times 1670}{38} = 9832.21$. $S_{xx} = 166991 - \frac{2436^2}{38} = 10830.58$.

Hence $\hat{\beta} = 0.9078$. And $\hat{\alpha} = 43.95 - 64.11 \times 0.9078 = -14.25$.

(ii) The substantial negative value for $\hat{\alpha}$, and the scatter plot, indicate that marks on SM are lower (harder to get) than on PS. There is a clear relation between the two, with $\hat{\beta}$ not far from 1, suggesting that the same types of skill and understanding are being examined in both.

At the upper end, the marks for SM are above the fitted line, which may just be due to having three very good students or it may perhaps suggest trying a curved relation to fit the whole data and tail off towards the origin.

(iii) The total sum of squares $S_{yy} = 86402 - \frac{1670^2}{38} = 13009.89$ (with 37 df).

The regression sum of squares is $\frac{{S_{xy}}^2}{S_{xx}} \left(\text{or } \hat{\beta}S_{xy} \text{ or } \hat{\beta}^2 S_{xx}\right) = 8925.68$ (with 1 df).

Hence the residual SS is 13009.89 – 8925.68 = 4084.21 with 36 df, and the residual mean square is 4084.21/36 = 113.45.

This (113.45) is the estimate ($\hat{\sigma}^2$) of σ^2.

$\text{Var}\left(\hat{\beta}\right) = \frac{\sigma^2}{S_{xx}}$, so the estimated variance of $\hat{\beta}$ is $\frac{\hat{\sigma}^2}{S_{xx}} = \frac{113.45}{10830.58} = 0.01047$.

$\text{Var}\left(\hat{\alpha}\right) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)$, so the estimated variance of $\hat{\alpha}$ is

$$\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right) = 113.45\left(\frac{1}{38} + \frac{64.11^2}{10830.58}\right) = 46.018.$$

Solution continued on next page

[NOTE. The test statistic for the usual F test for "the significance of the regression" is $\dfrac{8925.68}{113.45} = 78.68$, which is very highly significant as an observation from $F_{1,36}$.

This indicates that the line fits well.]

(iv) When $x = 80$, we have $\hat{y} = -14.25 + (0.9078 \times 80) = 58.37$.

The estimated variance is $\hat{\sigma}^2\left(\dfrac{1}{n} + \dfrac{\left(80 - \bar{x}^2\right)}{S_{xx}}\right) = 113.45\left(\dfrac{1}{38} + \dfrac{15.89^2}{10830.58}\right) = 5.63.$

Thus the 95% interval for $\hat{y}$, the mean mark at $x = 80$, is $58.37 \pm t_{37}\sqrt{5.63}$ where t_{37} represents the double-tailed 5% point of the t distribution with 37 df; we take this as (approximately) 2.02 here. Thus the interval is $58.37 \pm (2.02 \times 2.37)$, i.e. (53.6, 63.2).

This is an interval for the mean mark. Individual marks will vary around this [Var($\hat{y}_i$) has an extra term in it when interpreted as the estimated variance of an individual observation]. The scatter is quite large even though a line fits the data quite well.

Higher Certificate, Paper III, 2005. Question 4
[This solution continues on the next page]

(i) The measured response in a time series is often decomposed into trend, seasonal variation, cyclical variation and "irregular" or residual variation. The trend is a long-term change in the average response. Seasonal variation is a regular change which may occur quarterly, monthly, weekly, daily or on any other time-scale according to the nature of the series. Cyclical variation is likewise a regular change occurring (usually) on a much longer time-scale; this also depends on the nature of the series, but it will often refer to "economic cycles" or "business cycles" which are typically several years long and thus are qualitatively different from being seasonal. The irregular or residual variation may be genuinely random, or just irregular on a (much) shorter time-scale than the other components.

Using obvious notation, an additive model for the measured response Y_t at time t is

$$Y_t = T_t + S_t + C_t + I_t .$$

A multiplicative model is $Y_t = T_t S_t C_t I_t$ which becomes additive if logs (to any base) are taken: $\log Y_t = \log T_t + \log S_t + \log C_t + \log I_t$.

Notes.

(1) It can be difficult to identify cyclical components – a long series is often needed. A model might therefore omit the C_t component; if there is any cyclical variation present, it would then be bound up with the S_t (or possibly T_t) components.

(2) An alternative approach it to measure the trend as variation from an overall average X which then becomes a further component in the model.

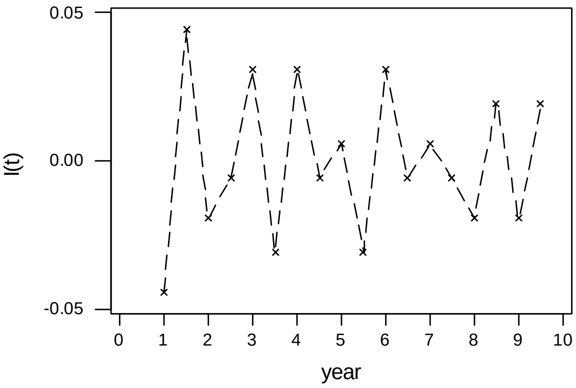
(ii) (a) There is an upward trend, possibly linear, but there is a suggestion of curvature of the type $y = \log x$. There is a strong seasonal “up–down” effect. Over this period of time it is not possible to say whether there is a cyclical component (as often happens with economic data). Besides the seasonal effect there may be some irregular variation.

(b) *MA1* refers to the moving averages at the first half-years and *MA2* to those at the second half-years. The first such figure is *MA2* at the second half-year of year 1. This is in the second cell of the table, so only three of the original observations can have gone into it. However $\frac{1}{3}(1.6+1.2+1.8)=1.533$ so this simple three-point average is not the form that has been used. Instead, a weighted moving average has been used, so as to reduce fluctuation in the detrended series. The weights are 1:2:1 over the three observations – we note that $1.45 = \frac{1}{4}\left((1\times1.6) + (2\times1.2) + (1\times1.8)\right)$. This can also be checked for the first *MA1* figure $[1.55 = \frac{1}{4}\left((1\times 1.2) + (2\times 1.8) + (1\times 1.4)\right)]$, and it works similarly for all the others.

A possible alternative weighted MA would use 4 observations with weights $\left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right)$. This would reduce fluctuation further, but would not be so appropriate for half-yearly data.

(c) The detrended data table is as shown on the left below. The entries in it are (observation – *MA*). Hence the seasonal components are +0.2056 and –0.2056, as shown at the foot of the table. These give the deseasonalised or irregular components as shown on the right.

			Deseasonalised	
Year 1	.	–0.250	.	–0.0444
2	0.250	–0.225	0.0444	–0.0194
3	0.200	–0.175	–0.0056	0.0306
4	0.175	–0.175	–0.0306	0.0306
5	0.200	–0.200	–0.0056	0.0056
6	0.175	–0.175	–0.0306	0.0306
7	0.200	–0.200	–0.0056	0.0056
8	0.200	–0.225	–0.0056	–0.0194
9	0.225	–0.225	0.0194	–0.0194
10	0.225	.	0.0194	.
	1.85	–1.85		
÷9 =	0.2056	–0.2056		



These deseasonalised or irregular components show no particular trend or pattern and so the additive model fits these data adequately.

<u>Higher Certificate, Paper III, 2005. Question 5</u>

(i) The number of defectives in a sample, X, has the binomial distribution with parameters 20 and p, i.e. $X \sim \mathrm{B}(20, p)$.

$$P(\text{accept batch} \mid p) = P(X = 0 \text{ or } 1 \mid p) = (1-p)^{20} + 20p(1-p)^{19}$$
$$= (1-p)^{19}(1 + 19p).$$

p = 0.01 : P(accept batch) = 0.9831
0.05 : 0.7358
0.1 : 0.3917

(ii) The batch is accepted in the following cases.

Number of defectives in first sample	
0	(Second sample not taken)
1	(Second sample not taken)
2	Second sample has 0 defectives

So P(accept batch)

$= (1-p)^{19}(1 + 19p)$
$+ P$(2 defectives in first sample <u>and</u> 0 defectives in second sample)

$$= (1-p)^{19}(1 + 19p) + \frac{20\times19}{2} p^2(1-p)^{18} \times (1-p)^{20}.$$

The values of this are as follows.

$p = 0.01$: 0.9831 + (0.01586×0.81791) = 0.9831 + 0.01297 = 0.9961

$p = 0.05$: 0.7358 + (0.18868×0.35849) = 0.7358 + 0.06764 = 0.8034

$p = 0.1$: 0.3917 + (0.28518×0.12158) = 0.3917 + 0.03467 = 0.4264.

Solution continued on next page

(iii)

<u>Scheme (i)</u> P(reject batch) = 0.0169 for $p = 0.01$
0.2642 for $p = 0.05$
0.6083 for $p = 0.1$.

Let S = total number inspected. $E(S) = 20P(\text{accept batch}) + 1000P(\text{reject batch})$.

The values of $E(S)$ are as follows.

$p = 0.01$: $(20\times0.9831) + (1000\times0.0169) = 36.6$

$p = 0.05$: $(20\times0.7358) + (1000\times0.2642) = 278.9$

$p = 0.1$: $(20\times0.3917) + (1000\times0.6083) = 616.1$.

<u>Scheme (ii)</u>

$E(S) = 20P(\text{accept batch based on first sample})$
$+ 40P(\text{2 defectives in first sample and 0 in second})$.

The values of $E(S)$ are as follows.

$p = 0.01$: $(20\times0.9831) + (40\times0.01297) + (1000\times0.0039) = 24.1$

$p = 0.05$: $(20\times0.7358) + (40\times0.06764) + (1000\times0.1966) = 214.0$

$p = 0.1$: $(20\times0.3917) + (40\times0.03467) + (1000\times0.5736) = 582.8$.

Scheme (ii) will have lower inspection cost than scheme (i) for these values of p.

Higher Certificate, Paper III, 2005. Question 6

(i) Simple random sampling, for samples of size n from a population of size N, is where every sample has the same probability of selection. (This probability is, of course, $1/\binom{N}{n}$.) A consequence of this is that every individual in the target population has the same probability of being selected for the sample.

If a population is not homogenous as a whole, but can be split into groups each of which is homogenous within itself, it will be better to select randomly within each group, i.e. stratified random sampling. This allows the different groups to be studied, as well as increasing precision of overall estimates. Also, when a very large population is to be sampled using, for example, a list of names, a systematic sample can be much easier to organise and may be treated as random provided any trends or cyclical patterns in the list are avoided.

(ii) (a) Errors in recording responses, due to poor training of enumerators or interviewers, and/or to carelessness or misunderstanding of subjects' answers. In a postal questionnaire, poor wording of questions may lead to respondents not answering the question intended.

(b) Transfer errors when data are taken from forms and entered into a processing system. Illegible answers could also occur on postal survey questionnaires.

(c) Non-response to postal surveys or refusal to co-operate/be interviewed. This may happen because of lack of interest in the topic being studied, objection to the wording of the questions or the approach of the interviewer, unwillingness to give time to answering, or simply being asked too often to take part in a survey.

(d) Failure to locate individuals/units chosen to take part in a survey. This may for example happen because of faulty lists, non-availability at the time an interviewer calls, premises being empty because people have moved, or different work and/or leisure habits so that individuals would need to be contacted at unusual times not planned for in the survey.

(iii) Telephone surveys only contact people available and willing to answer at the time of ringing, who have some interest in the topic under study, and whose numbers are not ex-directory (if a telephone directory is used as a sample frame). High rates of refusal to respond are likely from people who have been contacted frequently for such surveys. Further, in some countries by no means everyone has a telephone.

Solution continued on next page

(iv) Let n_1 = 1015, n_2 = 1005 be the numbers of people questioned in years 1, 2. Then $\hat{p}_1 = \frac{853}{1015} = 0.8404$ and $\hat{p}_2 = \frac{780}{1005} = 0.7761$ are the estimates of the proportions in favour.

If p_1, p_2 are the true proportions in the population, the null hypothesis is $p_1 = p_2$ and the alternative hypothesis is $p_1 > p_2$. We may use the Normal approximation to the binomial for these values of n and p. The test will be one-sided.

We have

$$\text{Var}\left(\hat{p}_1 - \hat{p}_2\right) = \frac{p_1\left(1-p_1\right)}{n_1} + \frac{p_2\left(1-p_2\right)}{n_2}$$

which is estimated by $\frac{0.8404 \times 0.1596}{1015} + \frac{0.7761 \times 0.2239}{1005}$ = 0.00030505.

Thus the value of the test statistic is

$$\frac{0.8404 - 0.7761}{\sqrt{0.00030505}} = \frac{0.0643}{0.0175} = 3.7 .$$

This is very highly significant as an observation from N(0, 1) [the upper single-tailed 0.5% point is 2.576].

There is strong evidence against the null hypothesis in favour of the alternative hypothesis which says that there <u>is</u> a decrease in the proportion of supporters.

Higher Certificate, Paper III, 2005. Question 7

(i) Let X denote the underlying random variable, with pdf $f(x) = \frac{1}{\mu} e^{-x/\mu}$. So the likelihood function is $L(x_1, ..., x_n) = \mu^{-n} \prod_{i=1}^{n} e^{-x_i/\mu}$ and $\log L = -n \log \mu - \sum_{i=1}^{n} x_i / \mu$.

$\therefore \frac{\partial(\log L)}{\partial \mu} = -\frac{n}{\mu} + \frac{1}{\mu^2} \sum x_i$, which we set equal to 0 to obtain the MLE $\hat{\mu}$. Hence

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad [= \bar{x}].$$

(We can quickly confirm that this is a maximum by considering the second derivative.)

(ii) The value of the estimate $\hat{\mu}$ is $\bar{x} = 2989.8$.

For any specified x, $P(X \le x) = \int_0^x \frac{1}{\mu} e^{-\frac{t}{\mu}} dt = \left[-e^{-\frac{t}{\mu}} \right]_0^x = 1 - e^{-\frac{x}{\mu}}$.

x	$\frac{x}{\hat{\mu}}$	$1 - e^{-\frac{x}{\hat{\mu}}}$	$\times 96$	
1000	0.33447	0.2843	27.29	} Expected
3000	1.00341	0.6334	60.80	} frequencies up
8000	2.67576	0.9311	89.39	} to the given
10000	3.34471	0.9647	92.61	} values of x

In the category $500 \le x < 1000$, the expected frequency is 27.29 – 14.78 = 12.51.

Hence up to 2000 the expected frequency is 46.82, and so in $2000 \le x < 3000$ it is 60.80 – 46.82 = 13.98.

Up to 6000, the expected frequency is 83.10. So in $6000 \le x < 8000$ the frequency is 89.39 – 83.10 = 6.29. Similarly, in $8000 \le x < 10000$ it is 92.61 – 89.39 = 3.22.

We should check that, with these values put in the table, the total expected frequency is 96. It is. It is usually argued that some grouping of cells is needed for chi-squared goodness-of-fit tests where there are "small" expected frequencies, "small" often being interpreted as < 5. We take the top group here as "all $\ge$ 8000". This gives

Solution continued on next page

Upper end of interval	250	500	1000	1500	2000	3000	4000	5000	8000	>8000
Observed frequency	11	16	16	10	10	11	7	5	4	6
Expected frequency	7.70	7.08	12.51	10.58	8.95	13.98	10.01	12.29	6.29	6.61

There are 10 cells in the table and one parameter has been estimated, so there will be 8 degrees of freedom for the chi-squared test.

The value of the test statistic is

$$X^2 = \frac{(11-7.70)^2}{7.70} + \frac{(16-7.08)^2}{7.08} + \ldots + \frac{(4-6.29)^2}{6.29} + \frac{(6-6.61)^2}{6.61} = 20.54 .$$

This is referred to χ_8^2. It is significant at slightly beyond the 1% level (critical point is 20.09).

Hence the (strong) evidence is that the model does not fit the data. The largest contributions to X^2 come from the first two intervals, especially the (250, 500) interval, and from the (4000, 6000) interval.

(iii)

$$P(X > 20000) = \int_{20000}^{\infty} \frac{1}{\mu} e^{-x/\mu}\, dx = \left[-e^{-x/\mu}\right]_{20000}^{\infty} = e^{-20000/\mu} .$$

Estimating this using $\hat{\mu} = \bar{x} = 2989.8$, we get 0.01244.

However, the data give that the relative frequency of claims above 20000 is 2/98 = 0.02083. The model substantially underestimates the probability of claims of this size.

(iv) The distribution of the number of claims is very skew, with a long tail to the right. Also, the frequency in the second interval is greater than in the first (of the same width), so a model needs to have a mode above 250. The rate of decrease of frequencies is slow. Perhaps a log-normal model might be better, or a more general 2-parameter model such as a gamma.

Higher Certificate, Paper III, 2005. Question 8

An obvious characteristic of the data is the difference between length and width. The "squares" are in every position rectangles, and the range of length sizes in any position (1 to 7) does not even overlap with the range of width sizes. Length is lower in positions 6 and 7 across the tray than it is elsewhere, and width is highest in position 7. Both length and width measurements are about equally variable in all positions.

Height is much more variable, especially in relation to its mean size. It also increases fairly steadily from position 1 to 7 (though 5 goes against this trend), with position 7 being particularly variable, due perhaps to one or two very large values (max 38).

There could be a temperature gradient in the oven, related to width, which affects height, and some other trends which result in length and width not being the same although the original material was presumably squarely placed.

If the appearance and uniformity of the "square" product are important, some attention needs to be given to the operation of the oven.

Summary of average (mean) and range (maximum – minimum in the whole data):

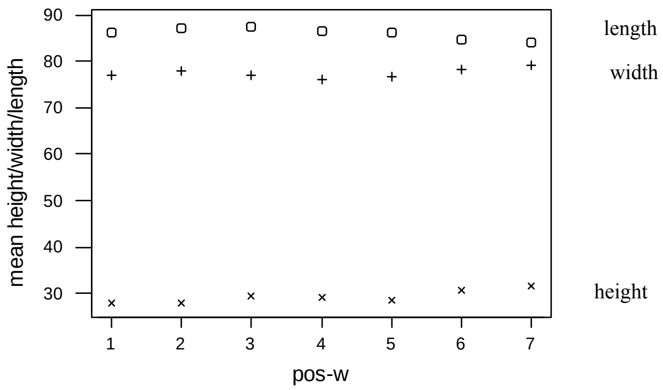
	Average	Range
Length	86.2	8
Width	77.5	11
Height	29.2	14

The combined effect on volume is to produce larger values in positions 6 and 7, with 1 to 5 showing an increase followed by a decrease. High variability is noted in 7, and fairly high in 2.

Data were collected just after removal from the oven. It is quite possible that after cooling some of the characteristics measured would have settled down more. We might usefully be told how many people were involved in measuring the data, as there could have been a time effect while collecting it.

Solution continued on next page

One useful diagram is to show the mean measurements against pos-w:



The above figure shows all three sets of data on the same scale. This is very useful in comparing length and width, but putting height on the same diagram hides the detail of the changes in the others because of the vertical scale. Two separate diagrams might be better.

The figure below shows an interesting comparison – volume depends quite closely on height. The numbers in brackets show positions across the width of the tray, 1 to 7.