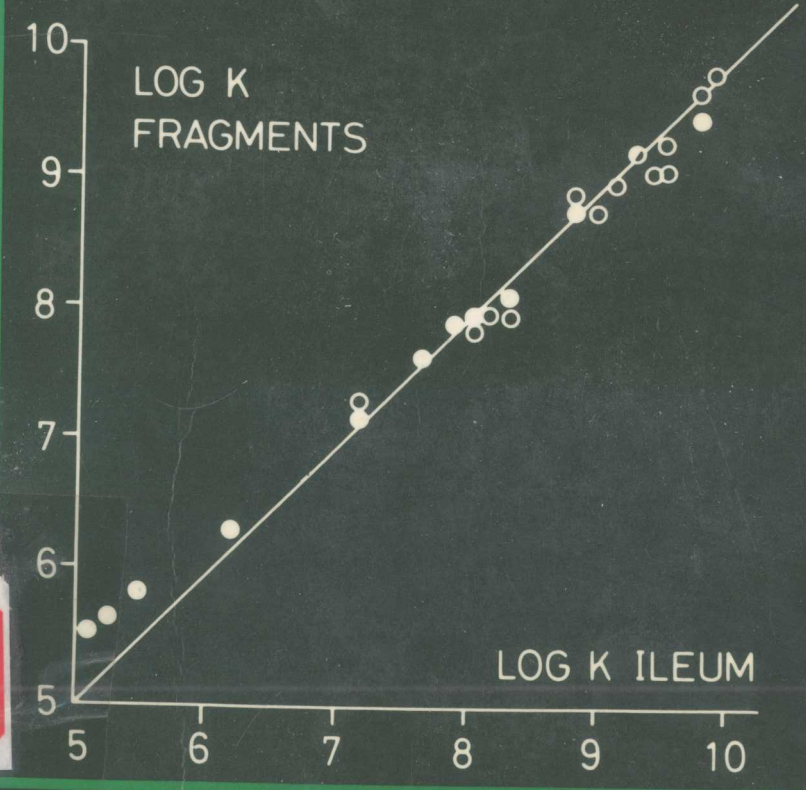
Quantitative Aspects of Chemical Pharmacology

R. B. Barlow



Quantitative Aspects of Chemical Pharmacology

CHEMICAL IDEAS IN DRUG ACTION WITH NUMERICAL EXAMPLES

R.B. BARLOW, MA, BSc., D Phil.

Lecturer in Chemical Pharmacology,
University of Bristol

CROOM HELM LONDON

Croom Helm Ltd, 2-10 St John's Road, London SW11
ISBN 0-85564-892-2

British Library Cataloguing in Publication Data

Barlow, R.B.
Quantitative aspects of chemical pharmacology.
1. Structure-activity relationship
(Pharmacology)
I. Title
615'.7 QP906.S75

0-85664-892-2
0-7099-0300-6(pb)

Reproduced from copy supplied
printed and bound in Great Britain
by Billing and Sons Limited
Guildford, London, Oxford, Worcester

Acknowledgements

I wish to thank the following for permission to include material which is copyright:
Dr R. C. Campbell and the Cambridge University Press for Tables 1.1, 1.2 and 1.3, which are shortened versions of Tables A.12, A.15 and 16, and A.3 in *Statistics for Biologists*; the Editors of the *Journal of Pharmacy and Pharmacology* for Table 1.4, which is taken from the paper by C. I. Bliss (1938), *Quart. J. Pharmac.*, no. 11, p. 195; Prof. B. Pullman and Academic Press for Figure 3.16 which is taken from the paper by J. L. Coubeils and B. Pullman (1972) *Mol. Pharmacol.*, no. 8, p. 281; the Editorial Board of the *British Journal of Pharmacology* for Figure 5.4 which is taken from the paper by F. B. Abramson, R. B. Barlow, M. G. Mustafa and R. P. Stephenson (1969) *Br. J. Pharmac.*, no. 37, p. 223.

I am most grateful to Dr P. S. Brown and Dr F. Roberts for valuable comments on Chapter 1 and to Dr Brown, Dr P. V. Taberner, Dr N. J. M. Birdsall and Prof. B. L. Ginsborg for material used in the examples. I thank Prof. J. F. Mitchell for his encouragement.

Finally I wish to thank Mrs Pat Berman for her expert typing and Mrs Anne Duncan for the illustrations.

I should be very glad to be informed of errors, omissions or misrepresentations.

R. B. Barlow

Department of Pharmacology
Medical School
Bristol.

Foreword

This book gathers together chemical ideas which are important for understanding how drugs act and how new drugs may be developed. Students meet some of these ideas in courses in chemistry, biochemistry and pharmacology but in my experience they find it difficult to put together information often acquired in different years and from different departments; they need it set out in a book.

I believe also that it helps if they can see how these ideas may be applied to numerical problems so I have included examples (with answers), many of which are taken from research work. These should therefore provide a chance for the student both to test his or her own understanding and to see the sort of results which form the bricks of which any experimental science is made. In all the problems the arithmetic takes less than ten minutes with a small calculator which has exponential functions and logarithms.

From this collection of information and ideas the book proceeds to consider some of the attempts which have been made to relate chemical properties quantitatively to biological activity and to predict the activity of new compounds from results obtained with a carefully selected trial group. It is important that the medicinal chemist and the biologist work together in this aspect of drug development and the book attempts to explain what the medicinal chemist is doing and to consider some of the assumptions he makes.

Contents

Acknowledgements

Foreword

1. Measuring Drug Activity 1

Introduction 1

Measuring the Effects of Drugs: 'Responses' 2

Assessing Biological Activity 14

Application of Bioassay Methods to the Comparison of Drugs 29

Types of Drug Action 34

2. Physicochemical Problems 39

Rates of Reaction and Positions of Equilibrium 39

Solubility 46

Volume of Dilution 48

Rate of Removal 49

Rate of Absorption 52

Partition Coefficients and Crossing Membranes 60

Relations between Reaction Rates and Equilibrium Constants 70

3. The Size and Shape of Molecules 80

Size: Molal Volumes and Apparent Molal Volumes 80

Shape 83

Isomerism 85

Optical Rotation 87

Stereochemical Purity and Biological Activity 88
Description of the Arrangements of Groups about an Asymmetric Carbon Atom 93
Preferred Conformations 107

4. Chemical Processes where Size and Shape are Important 117
Adsorption 117
Adsorption of Two Species 129
Enzyme and Substrate 137
The Need for Other Models 150
Agonists and Receptors 151
Rate Theory 170
Desensitisation of Receptors 173
Forces between Drug and Receptor 181

5. The Prediction of Drug Activity 187
Nonspecific Activity 187
Homologous Series 191
Biological Activity and Chemical Properties: the Ideas of Hansch 196
Is Affinity Predictable? 203
Conclusion 220

Appendix: Line-fitting by the method of least-squares 224
Answers to Examples 232
Index 246

References are numbered and listed at the end of each chapter. The Tables for chapter 1 contain statistical information and are collected together on pages 36–8. The other Tables and all the Figures are inserted into the text where appropriate.

CHAPTER 1

Measuring Drug Activity

'A branch of science comes of age when it becomes quantitative.'
J. H. Gaddum, *Edinburgh Medical Journal* (1942) no. 49, p. 731.

Introduction

The development of drugs starts with the recognition that some particular pharmacological effect may be useful therapeutically. Sometimes an application can be seen before there is a compound with the desired properties. Past history often suggests the reverse: several important drugs, such as the sulphonamides, have been made by the chemist years before their therapeutic value was realised. In either situation an initial discovery depends upon both an understanding of what may be useful and upon a knowledge of what drugs do; it requires both the flash of genius and the humdrum collection of pharmacological information.

Once something potentially useful has been found it is necessary to know not only what a substance does but how well it does it, so that comparisons can be made and better drugs discovered. This may not be easy. The effects which the pharmacologist is called upon to study are extremely diverse. Some are much more difficult to detect and compare than others and less reliance can be placed in the results. The effects of a drug on the growth or function of cells may be relatively easy to see or to measure and should be highly reproducible; effects on the central nervous system resulting in changes in behaviour are likely to be a very different matter. Drug development depends critically upon adequate testing and this demands an appreciation of the problems of quantitative pharmacology.

When a suitable pharmacological test has been devised the successful development of new drugs depends upon finding what changes in chemical structure lead to increased activity. It may be possible from the relation between structure and activity to obtain some information about how the drugs act, which in turn may suggest new structures which are worth investigation. This part of

the work is both a scientific problem in chemical ('molecular') pharmacology and of great practical importance, because the making and testing of many new compounds is expensive and takes time. It is an intellectual challenge to know why activity alters with chemical structure but the answers may also make it possible to develop better drugs rapidly and economically.

Because drug development is impossible without proper quantitative pharmacological tests the book begins with a consideration of the problems of measuring biological activity.

Measuring the Effects of Drugs: 'Responses'

If a drug produces an effect, it changes the properties of some of the cells in an organism. Some effects may be actually visible but most can only be detected with suitable apparatus. Many effects involve physical or chemical changes which can be measured. These can be classified into:

(i) changes in mechanical properties, e.g. in length, volume, pressure, flow, rates of contraction;
(ii) changes in electrical properties, e.g. in the polarisation or conductance of the cell membrane, potential changes due to the activity of the heart or the central nervous system, changes in rates of firing of neurones;
(iii) chemical changes, e.g. in the concentration of substances involved in cell function, such as nervous transmission, or in metabolism; this may include changes in the amounts of enzymes involved in the synthesis or metabolism of such substances.

Although the apparatus used to detect such effects may be sophisticated, the response is usually the simplest type to consider, because it can be measured and its size depends, within limits, on the dose.

In some tests the effect cannot be measured, perhaps because it is too complex, but it may be possible to devise some scoring system which allows effects to be compared and arranged in order. This often applies in tests for the effects of drugs on behaviour and in the clinical assessment of the treatment of disease.

In other situations, such as toxicity tests, the responses can only be 'all-or-none'. An animal or a cell either lives or dies; a nerve cell either conducts or it does not; a striated muscle cell either twitches

or it does not. Responses can be made to appear graded by taking groups of animals or cells or working with a nerve trunk or a muscle bundle, rather than with single cells. Because individuals differ in their sensitivity to a drug there will be a dose range within which only a proportion will be affected, rather than all of them or none at all. The response of a group in this situation is termed quantal. The effect depends on the proportion of units responding in an all-or-none fashion and it varies with the dose because of the differences in the sensitivity of individual units. This is not the same as where the performance of the unit is variable and the response is termed 'graded'. For example effects involving smooth muscle are graded but those involving twitches from striated muscle are quantal. Effects can be classified into:

(i) what can be measured and gives a graded response;
(ii) what can be scored and ranked;
(iii) what can be counted or depends on the proportion of units responding in an all-or-none fashion.

The complexity of the apparatus used to detect the effect is no guide to the complexity of a test. Elaborate equipment is usually employed in order to simplify the response but this may also be simplified by a suitable choice of test preparation. By working with enzymes, single cells, or even pieces of tissue, rather than whole animals it should be possible to simplify the nature of the response by reducing the number of variables on which it depends. In particular it should be possible to reduce, or even eliminate, the complicating effects of time on the response. In a whole animal the effect of a drug usually rises to a peak (depending on how it is administered) and then declines. What happens at any particular moment depends on:

(i) how the drug reaches its site of action;
(ii) what it does when it gets there;
(iii) how it is removed.

These processes may well be quite unrelated to each other. In a test on an isolated preparation with the drug added directly to the bathing fluid it may be possible to see what a drug does without the complications of transport and removal. In some situations it is even possible to make measurements in conditions of equilibrium,

or at least in a steady state, and so to measure effects at a much more fundamental level. Such tests, however, are far removed from the clinical situation and there is therefore a conflict between the need to simplify to obtain reliable tests and the difficulty of simplifying tests for properties which are clinically most interesting. In fact a single type of test is not enough; all three processes, transport, action and removal, need separate investigation if the effects in the whole animal or man are to be understood.

Reproductibility of Graded Responses

The biggest problem in all drug testing is the way the response may vary from one experiment to another. It is not enough that the effect of a dose of one compound appears to be bigger than that of a dose of another. It is necessary to ask whether there really is any difference between the two results and to see if the difference is big enough to refute the null hypothesis, that the results could all have come from the same group.

If the test preparation produces a graded response and the variation in response is due only to chance, individual responses should be distributed *normally* about the true value. The distribution is Gaussian and can be seen in Figure 1.1, where the probability, P, of obtaining a result is plotted against its distance from the mean, μ, expressed as a multiple, N, of the standard deviation, σ.

If the same dose has been tested n times and the individual responses are $r_1, r_2, r_3, \ldots, r_n$, the *mean*, $\bar{r} = \mathrm{S}r/n$ where $\mathrm{S}r = r_1 + r_2 + r_3 + \ldots + r_n$. An estimate of the *variance* of the results, $V = \mathrm{S}d^2/(n-1)$, where $\mathrm{S}d^2$ is the sum of the squares of the deviations from the mean, i.e. $\mathrm{S}d^2 = \mathrm{S}(r-\bar{r})^2$. There is actually no need to calculate the deviation of each individual result from the mean because $\mathrm{S}r^2 = \mathrm{S}(\bar{r}+d)^2 = n\bar{r}^2 + 2\bar{r}\,\mathrm{S}d + \mathrm{S}d^2$, but $\mathrm{S}d = 0$, so $\mathrm{S}d^2 = \mathrm{S}r^2 - n\bar{r}^2 = \mathrm{S}r^2 - (\mathrm{S}r)^2/n$.

It is therefore only necessary to sum r and r^2, a simple operation for a calculator, and the mean and the variance of the results can be worked out. An estimate of the *standard deviation* of the results, $s = V^{1/2}$. The mean, $\bar{r}$, is an estimate of the true mean, μ; the standard deviation, s, is an estimate of the true standard deviation, σ.

For example if there were six responses: 15, 19, 21, 23, 18, 20, the mean, $\bar{r} = 116/6 = 19.3$; $\mathrm{S}r^2 = 2280$, $\mathrm{S}d^2 = 2280 - 6(19.3)^2 = 2280 - 2242.7 = 37.3$: $V = 7.5$ and $s = 2.7$.

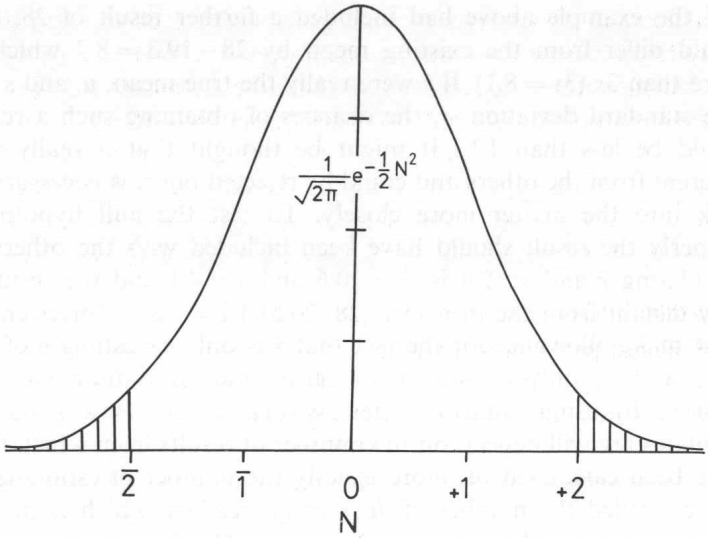


FIGURE 1.1 The graph shows the function

$$\frac{1}{(2\pi)^{1/2}}e^{-\frac{1}{2}N^2},$$

i.e. dP/dN, plotted against N. The probability

$$P = \int_{N_1}^{N_2}\left(\frac{1}{(2\pi)^{1/2}}e^{-\frac{1}{2}N^2}\right)dN$$

and the integral between $N = -1$ and $N = +1$ encloses 68.2% of the total area, i.e. there are roughly two chances in three that a result will lie within the range $\mu-\sigma$ and $\mu+\sigma$. The integral between $N = -2$ and $N = +2$ encloses 95.5% of the total area so there is only roughly one chance in twenty that a result will lie outside the range $\mu-2\sigma$ and $\mu+2\sigma$, which represent the two 'tails' (shaded). For a single tail there is only one chance in forty that a result will lie above $\mu+1.96\sigma$ and one in forty that it will lie below $\mu-1.96\sigma$.

If the individual results are distributed normally 68.2% of the values should lie within the range $\mu \pm \sigma$ and 95.5% should lie within the range $\mu \pm 2\sigma$. There is therefore less than one chance in twenty (5%) of obtaining a value outside this range. In the above example two of the results (15, 23) lie outside the range $\bar{r} \pm s$; none lies outside the range $\bar{r} \pm 2s$ (but note that $\bar{r}$ is only an estimate of μ and s is only an estimate of σ).

Our knowledge of the properties of the normal distribution makes it possible to assess the chances of obtaining a particular result, or group of results, and so perhaps to refute the null hypothesis by showing that the chances that the results all come from the same group are too small to be acceptable.

If the example above had included a further result of 28, this would differ from the existing mean by $28-19.3=8.7$ which is more than $3s$ ($3s=8.1$). If $\bar{r}$ were really the true mean, μ, and s the true standard deviation, σ, the chances of obtaining such a result would be less than 1%. It might be thought that it really was different from the others and could be rejected but it is necessary to look into the matter more closely. To test the null hypothesis properly the result should have been included with the others in calculating $\bar{r}$ and s; if it is, $\bar{r}=20.6$ and $s=4.1$ and the result is only distant from the mean by $(28-20.6)/4.1=1.8s$. Moreover, we must make allowance for the fact that $\bar{r}$ is only an estimate of the mean and s is only an estimate of the standard deviation. This can be done by using 'student's t-deviate' (Table 1.1). How good the estimates are will depend on the number of results from which they have been calculated or, more exactly the number of estimates of scatter, called the number of *degrees of freedom*, which is $(n-1)$. (With a mean value and $(n-1)$ results, the last result can be calculated.) On the normal curve 95% of the results lie within the range mean ± 1.96 times the standard deviation (95.5% within $\pm 2\sigma$), but if we are using estimates instead of true values, the number must be replaced by a bigger value, t, whose size depends on the number of degrees of freedom. In the example there are six degrees of freedom and the value of t is 2.45. The expectation is, therefore, that 95% of the results lie within the range $20.6 \pm 2.45 \times 4.1$, i.e. between 10.6 and 30.6, so the extra result (28) could well belong to the same group as the others. It may be different from the others but the evidence is not convincing.

The choice of probability which is considered acceptable depends upon the circumstances. If the probability of obtaining a result is less than one in twenty ($P<0.05$) it is often regarded as being significantly different but there may be situations where it is desirable to set the limit at one in one hundred ($P<0.01$). On the normal curve this includes all values in the range $\mu \pm 2.58\,\sigma$; values of t for this level of probability are included in Table 1.1.*

To compare different drugs it is clearly necessary to work with groups of results: single results are of little value. For instance the original six responses, with a mean of 19.3 and a standard deviation of 2.7, might have been obtained with a dose of one drug and in the same conditions a dose of another drug might have produced responses which were: 25, 26, 32, 29, 25, 31 (mean = 28; $s=3.1$).

* The statistical Tables 1.1 to 1.4 will be found at the end of this chapter, on pp. 36–8.

Could these all belong to the same group? The mean values differ by $28-19.3=8.7$: at what level of probability would this difference be significant?

If individual responses are distributed normally, the means of groups of results will also be distributed normally though they will be gathered much more closely around the true value. The standard deviation of the distribution of means is estimated by the *standard error* of a group of results, $se = s/n^{1/2}$. The values for the two sets of results are $se_1 = 2.7/6^{1/2} = 1.1$ and $se_2 = 3.1/6^{1/2} = 1.3$. If two means have true standard deviations σ_1 and σ_2, the standard deviation of the difference between two means is:

$$\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)^{1/2}$$

where n_1 and n_2 are the numbers of responses from which the means are calculated. An estimate of this standard deviation is:

$$\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^{1/2} = (se_1^2+se_2^2)^{1/2} = (1.1^2+1.3^2)^{1/2} = 1.7.$$

The difference between the means is therefore $8.7/1.7 = 5.1$ times the estimate of the standard deviation. There are five estimates of scatter from each group, making ten degrees of freedom in all. The corresponding value of t with $P = 0.05$ is 2.23 and with $P = 0.01$, t is 3.17. The difference is therefore more than would be expected even at the '1 %' level of probability.

The example given above shows the value of working with groups of responses. If only three responses were obtained in the second group, 25, 25, 26, the difference between the means is $25.3-19.3 = 6.0$. The standard error of the second group is 0.3 and the estimate of the standard deviation of the difference is $(1.1^2+0.3^2)^{1/2} = 1.15$ and the difference between the means is $6.0/1.15 = 5.2$ times the estimate of the standard deviation. With seven degrees of freedom the value of t with $P = 0.05$ is 2.36 and with $P = 0.01$ t is 3.50 so the difference is significant even at the more rigorous level. Although the difference between the means is less than with all six responses in the second group, the standard error is now very small. Indeed the scatter appears suspiciously different in the two groups, though the experiment is poorly designed for testing this because there are only three responses in the second group compared with six in the first.

To investigate the matter, an analysis of variance can be made.

For two groups, size n_1 and n_2, with means $\bar{r}_1$ and $\bar{r}_2$ and an overall mean $\bar{r}$,

$$S(r-\bar{r})^2 = S(r_1-\bar{r}_1)^2 + S(r_2-\bar{r}_2)^2 + n_1(\bar{r}_1-\bar{r})^2 + n_2(\bar{r}_2-\bar{r})^2.$$

(These can easily be calculated from values of Sr and Sr^2, see above.) The first two terms estimate the variance within the groups; the last two represent the variance between the two means. If they are not very different $\bar{r}_1$, $\bar{r}_2$ and $\bar{r}$ will be close together and these last two terms will be small. For the two groups, 15, 19, 21, 23, 18, 20 and 25, 25, 26, the actual figures are:

$$110.1 = 37.3 + 0.7 + 6(19.3-21.3)^2 + 3(25.3-21.3)^2$$
$$= 38.1 + 24 + 48 \ (= 110.1)$$

This information can conveniently be arranged:

	ss	*df*	*V*
Between groups	72.0	1	72.0
Within group 1	37.3	5	7.46
group 2	0.7	2	0.35
Total	110.1	8	13.76

(where *ss* indicates the sums of the squares of the deviations, *df* is the number of degrees of freedom and *V* is the estimate of the variance. Note that for the variance between two means there is only one degree of freedom).

The variance within the first group is much bigger than the variance within the second and the ratio is $7.48/0.35 = 21.4$. It is possible to calculate the limiting value of this variance ratio, F, consistent with both sets of values having the same distribution. With $P = 0.05$, and the greater estimate based on five degrees of freedom and the lesser estimate based on two degrees of freedom, the limiting value of F is 39.3 (Table 1.2A), so it is not possible to confirm the suspicion. Table 1.2A is based on a two-tailed distribution, where no assumptions are made about which may be the greater variance. If it is known from the experimental plan that one should be bigger than the other, the test involves a one-tailed distribution (Table 1.2B) and the ratio would be significant with $P = 0.05$ (F is 19.3). This is not known, however, but even so it is doubtful whether we can assume that the variances are the same and can be pooled and used in a t-test.

If the group of three results with the second compound is

replaced by the original group of six values, 25, 26, 32, 29, 25, 31, the overall mean $\bar{r} = 23.7$, $\bar{r}_1 = 19.3$, $\bar{r}_2 = 28.0$, and the sums of squares of the deviations (*ss*), number of degrees of freedom (*df*), estimates of the variance (*V*) and the variance ratios (*F*) are:

	ss	*df*	*V*	*F*
First group	37.3	5	7.5	3.76
Second group	48.0	5	9.6	2.95
Total	310.5	11	28.2	

The values of *F* indicate that the variance in either of the two groups cannot be distinguished from that of all the results taken together. The total 'within sample' variance is $(37.3+48.0)/10 = 8.5$. The 'between sample' variance is $n_1(\bar{r}_1-\bar{r})^2+n_2(\bar{r}_2-\bar{r})^2$ divided by the number of degrees of freedom, which is only 1 because we are considering the means of the two groups, so this is $6\times 4.3^2+6\times 4.3^2 = 221.9$. We have therefore:

		df	*F*
Between sample variance	221.9	1	
Within sample variance	8.5	10	$\frac{221.9}{8.5} = 26.1$

With one degree of freedom for the greater estimate and ten degrees of freedom for the lesser estimate the limiting value of *F*, $P = 0.05$ is 6.94 (Table 1.2A) so the analysis of variance establishes that the difference between the means is greater than would be expected on the null hypothesis at this level of probability. If it is known that the results in one group should be bigger than those in the other, for instance if they were obtained with the same compound but with a larger dose, the test involves a one-tailed distribution and the variance ratio is significant even with $P = 0.01$ (Table 1.2B).

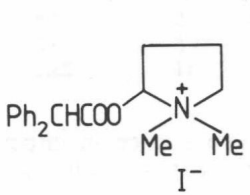
An analysis of variance would be expected to confirm the findings of a *t*-test but it is more rigorous, because it checks the assumptions that the sample variance is indistinguishable in both groups.

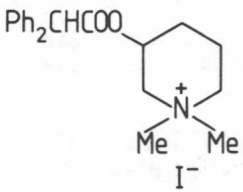
EXAMPLE 1:1.

Convince yourself that $Sd^2 = Sr^2 - n\bar{r}^2$ by checking, calculating the values for $r = 3, 4, 5, 6, 7$.

EXAMPLE 1:2.

The following estimates of log affinity constant (see page 158) for acetylcholine receptors in the guinea-pig ileum were obtained:





6.974	7.090
6.984	7.083
6.999	7.093
6.986	7.072
6.985	7.117
7.004	7.107
7.003	

Is there any evidence that the compounds have different affinities?

(From results of Abramson, PhD thesis, Edinburgh University, 1964.)

Ranked Responses

If the responses are not graded but can only be ranked, there is no reason to assume that they will be normally distributed. For instance, the steps in the scale used for scoring may be unequal. If the scores for two drugs were respectively 17, 18, 20, 21, 24 and 20, 23, 25, 26, 28, the values for the middle member of each set, called the *median*, are 20 and 25: 60% of the scores lie between 18 and 21 for the first group and between 23 and 26 for the second. The *interquartile* range, which gives the limits within which half the results lie is sometimes quoted as an indication of the scatter, but these particular results cannot satisfactorily be divided into quarters.

Could these results all belong to the same group? If they are arranged in rank order with the largest $=1$, the values become: 10,