

**Part III**  
**Appendices**

---

# Appendix A

## Exponentials and logarithms

---

Raising 10 to different powers is a familiar operation. For example,

$$10^1 = 10, 10^2 = 100, 10^3 = 1000, \dots$$

Mathematically this is regarded as a rule for getting from the power (1, 2, 3, etc.) to the value of 10 raised to that power (10, 100, 1000, etc.). The power is often referred to as the *exponent* and 10 raised to a power is called an *exponential* with base 10.

Raising 10 to a power can be extended to cover fractional powers using the convention that  $10^{\frac{1}{2}}$  stands for the square root of 10,  $10^{\frac{1}{3}}$  stands for the cube root of 10, and so on. The rule can also be extended to cover negative powers using the convention that  $10^{-1}$  stands for  $1/10 = 0.1$ . Table A.1 shows the rule for obtaining  $10^x$  from  $x$  for a variety of values of  $x$ .

Now suppose that we wish to go the other way and, starting with a value of  $10^x$ , find the value of  $x$ . For example, starting with 1000 gives  $x = 3$ , while starting with 0.1 gives  $x = -1$ . Starting with any positive number  $y$ , the value of  $x$  which makes  $10^x = y$  is called the *logarithm* of  $y$  with the base 10 and is written  $\log_{10}(y)$ . Taking logarithms with base 10 is the inverse operation to exponentiation with base 10. Thus  $10^3 = 1000$  and  $\log_{10}(1000) = 3$ .

**Table A.1.** Rules for finding  $10^x$  from  $x$

$x$	$y = 10^x$
0	1
1	10
2	100
3	1000
-1	0.1
-2	0.01
-3	0.001
$\frac{1}{2}$	$\sqrt{10}$
$\frac{1}{3}$	$\sqrt[3]{10}$

**Table A.2.** Multiplication using logarithms

Number		Logarithm
7.2	→	0.8573
16.9	→	1.2279
121.7	←	2.0852

Logarithms were introduced as a computational device in the seventeenth century to avoid multiplication and division. Tables were prepared so that the logarithm of any number could be looked up. Similarly, tables of exponentials were prepared so that logarithms could be converted back to the original numbers. These tables of exponentials were called *antilogarithms*. The use of logarithms to multiply 7.2 by 16.9 is shown in Table A.2. Arrows from left to right refer to looking up logarithms while arrows from right to left refer to looking up antilogarithms (exponentiation). The result line follows from addition on the logarithmic (right-hand) side or multiplication on the exponential (left-hand) side. The widespread availability of cheap electronic calculators means that nobody now uses logarithms for multiplication or division. However, their mathematical property of converting multiplication to addition, embodied in

$$\log(7.2 \times 16.9) = \log(7.2) + \log(16.9)$$

is still very useful. Another useful property which follows from this is that

$$\log(7.2^2) = 2 \times \log(7.2)$$

$$\log(7.2^3) = 3 \times \log(7.2)$$

and so on.

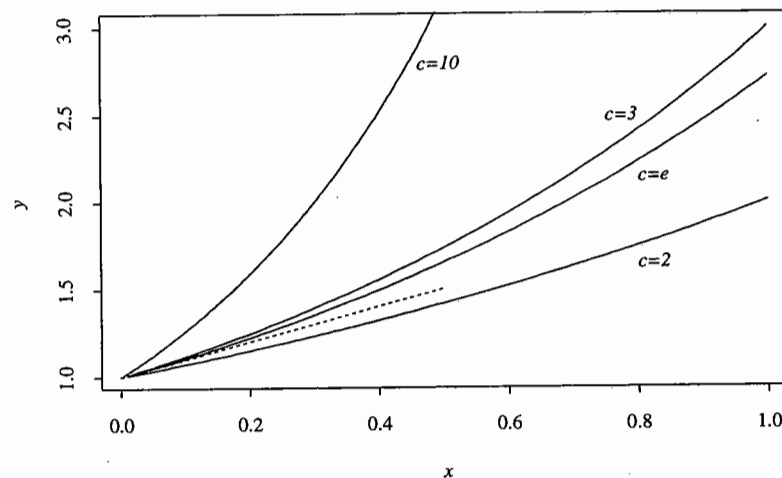
Raising 2 to a power is called exponentiation with base 2. The inverse process produces logarithms to the base 2 and these are written  $\log_2(y)$ . Both exponentials and logarithms can be defined with respect to any base. Fig. A.1 shows plots of the exponential functions  $10^x$ ,  $3^x$ ,  $e^x$ , and  $2^x$ , where the symbol  $e$  represents the number 2.71828183. The number  $e$  is chosen so that the tangent to the plot of  $e^x$  versus  $x$  drawn at  $x = 0$  has a slope of exactly 1 (shown by the broken line). It follows that *when  $x$  is very small*,

$$e^x \approx 1 + x.$$

and, therefore,

$$\log_e(1 + x) \approx x.$$

Logarithms to the base  $e$  are referred to as *natural* logarithms, and it is the above property that makes them 'natural'. The natural logarithm

**Fig. A.1.** Plots of the function  $y = c^x$ 

function is sometimes written as  $\ln(y)$ , but in this book we shall *always* use logarithms to the base  $e$ , and write them simply as  $\log(y)$ . We also write the exponential function with base  $e$  as  $\exp(x)$ . Note, however, that many electronic calculators assign an entirely different meaning to a key marked *exp*.

The logarithms of the same number, using different bases, are related by a simple constant multiplier. For example

$$\log_e(y) = \log_{10}(y) \times 2.3026$$

where  $2.3026 = \log_e(10)$ . Similarly

$$\log_2(y) = \log_{10}(y) \times 3.3219$$

where  $3.3219 = \log_2(10)$ .

## Appendix B

### ★ Some basic calculus

The *gradient* of the graph of  $y$  versus  $x$  measures the rate at which  $y$  is increasing (or decreasing) at any point on the graph. It is most easily defined for a straight line graph, such as the one in Fig. B.1. In this case the rate of increase or decrease is the same at any point on the graph, and is measured by the ratio of the *rise* to the *run*. For a straight line relationship in which  $y$  decreases with  $x$  the gradient is negative. Gradients have units equal to those of  $y/x$ . The central idea of calculus is that over a small run any curve is approximately a straight line and the gradient of the curve at any point in the run is approximately equal to the gradient of this line.

Differential calculus consists of a number of simple rules which are used to evaluate gradients of curves for which the  $y$  co-ordinate of any point on the curve is given by some function of the  $x$  co-ordinate. The most useful of these are shown in Table B.1. A further very important rule is that the gradient of a function constructed as the *sum* of two simpler functions is

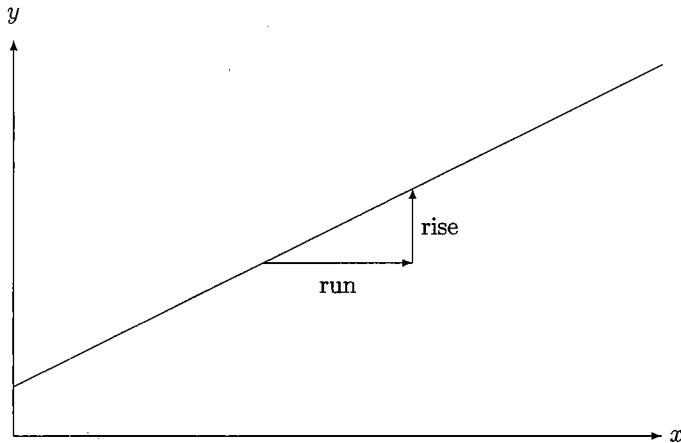


Fig. B.1. Gradient of a straight line graph

Table B.1. Gradients of some simple functions of  $x$

Function	Gradient
$c$ (constant)	0
$x$	1
$-x$	-1
$cx$	$c$
$(x)^2$	$2x$
$(x)^m$	$m(x)^{m-1}$
$\frac{1}{x} = (x)^{-1}$	$-(x)^{-2} = -\frac{1}{(x)^2}$
$\exp(x)$	$\exp(x)$
$\log(x)$	$\frac{1}{x}$
$(c+x)^2$	$2(c+x)$
$(c-x)^2$	$-2(c-x)$
$\log(c+x)$	$\frac{1}{c+x}$
$\log(c-x)$	$-\frac{1}{c-x}$

the sum of the gradients of the constituent functions so that, for example, the gradient of  $x + \log(x)$  is  $1 + 1/x$ .

The use of these rules is now illustrated by finding the gradient of the log likelihood for a rate  $\lambda$ , based on  $D$  cases and  $Y$  person years. The log likelihood for  $\lambda$  is

$$D \log(\lambda) - \lambda Y.$$

From Table B.1 the gradient of  $\log(\lambda)$  is  $1/\lambda$  and the gradient of  $\lambda$  is 1. Hence the gradient of the log likelihood is

$$\frac{D}{\lambda} - Y.$$

The maximum value of the log likelihood occurs when the gradient is zero, that is, when  $\lambda = D/Y$ , so the most likely value of  $\lambda$  is  $D/Y$ .

The curvature of the log likelihood curve at the peak is important in determining the range of supported values. A highly curved peak corresponds to a narrow range. The curvature at a point on a curve is a measure of how fast the gradient is changing from one value of  $x$  to the next; if the gradient is changing quickly then the curvature is high, while if the gradient is changing slowly the curvature is low. For log likelihood curves the gradient changes from a positive quantity (on the left) to a negative quantity (on the right) so the gradient decreases as  $x$  increases and the curvature is negative.

The curvature of a curve, at a point, is defined to be the rate of change of the gradient of the curve at that point. The way that Table B.1 can be used to find curvature is now illustrated using the log likelihood for  $\lambda$

again. The gradient of the log likelihood at any value of  $\lambda$  has been shown to be

$$\frac{D}{\lambda} - Y.$$

From Table B.1 the gradient of a constant is zero and the gradient of  $1/\lambda$  is  $-1/(\lambda)^2$ , so the curvature of the log likelihood at any value of  $\lambda$  is

$$-\frac{D}{(\lambda)^2}.$$

---

## Appendix C

### Approximate profile likelihoods

---



This appendix describes the mathematics underlying Gaussian approximation of profile log likelihoods.

#### C.1 The difference between two parameters

We shall start with an important special case. Consider a model with two parameters,  $\beta_1$  and  $\beta_0$ , and suppose that our main interest is in the *difference*

$$\gamma = \beta_1 - \beta_0.$$

We shall further assume that the log likelihoods for  $\beta_1$  and  $\beta_0$  are based on two independent sets of data so that the total log likelihood is the sum of the two separate log likelihoods.

Fig. C.1 illustrates the construction of the profile likelihood for  $\gamma$ . The upper panel of the figure shows the total log likelihood obtained by adding the log likelihoods for  $\beta_1$  and  $\beta_0$ . Contours are shown for log likelihood ratios of  $-5, -4, \dots, -1$ . The four diagonal lines correspond to different values of  $\gamma$ . For example, the top leftmost line represents values of  $\beta_1, \beta_0$  satisfying

$$\beta_1 - \beta_0 = 0$$

so that this line corresponds to  $\gamma = 0$ . Similarly, the remaining lines correspond to values of  $\gamma$  of 0.5, 1.0, and 1.5 respectively. To find the profile likelihood for  $\gamma$ , we find the maximum value of the log likelihood along each of these lines. This maximum is plotted against  $\gamma$  in the lower panel of the figure.

The Gaussian approximation of the profile log likelihood can be obtained from making use of the relationship between gradients and curvatures of the total log likelihood (upper panel), and the gradient and curvature of the profile log likelihood (lower panel). These relationships can be derived using the laws of calculus but are only quoted here.

If, at the maximum of the log likelihood along the line  $\beta_1 - \beta_0 = \gamma$ , the gradient is  $G_1$  with respect to  $\beta_1$  and  $G_0$  with respect to  $\beta_0$  the gradient

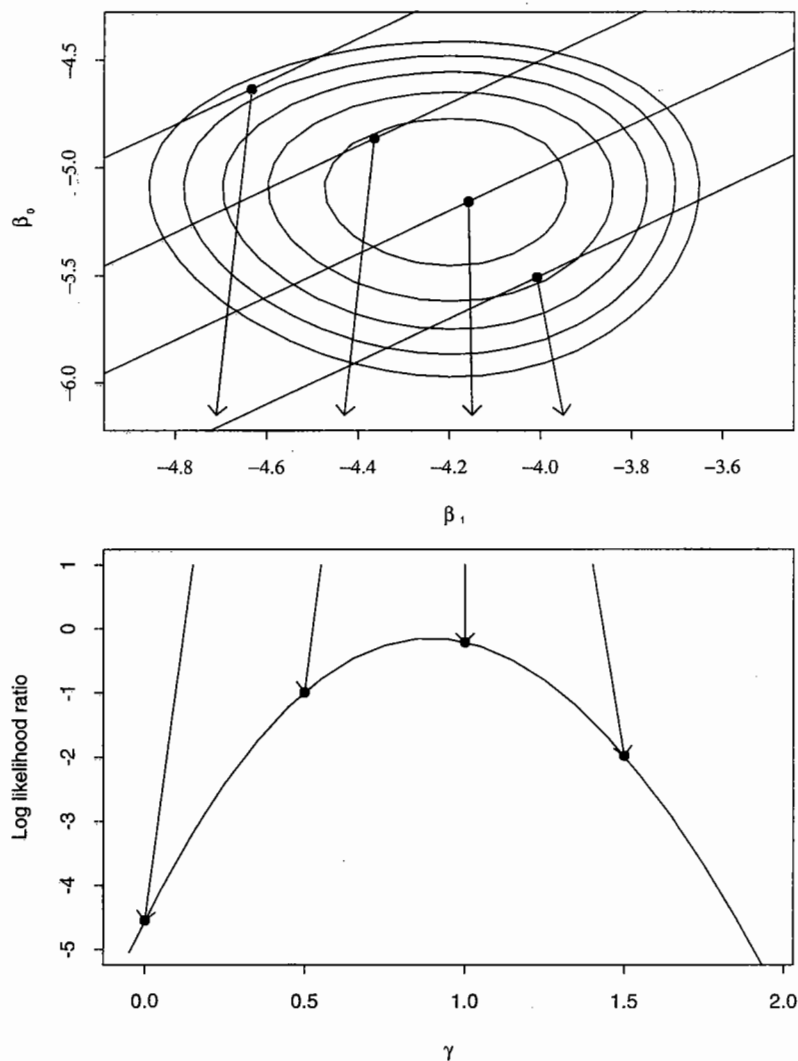


Fig. C.1. The profile log likelihood

of the profile log likelihood at  $\gamma$  is  $G$ , where

$$G = G_1 = -G_0.$$

If  $C_1, C_0$  are the corresponding curvatures with respect to  $\beta_1$  and  $\beta_0$ , then the curvature of the profile log likelihood at  $\gamma$  is  $C$ , where

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_0}.$$

From these results it follows directly that, if the most likely values of  $\beta_1$  and  $\beta_0$  are  $M_1$  and  $M_0$  respectively, and the corresponding standard deviations of the estimates are  $S_1$  and  $S_0$ , then the most likely value of  $\gamma$  is

$$M = M_1 - M_0,$$

and the standard deviation of the estimate is

$$S = \sqrt{(S_1)^2 + (S_0)^2}.$$

#### THE RATE RATIO REVISITED

As an example, we shall apply use these general rules to the problem of estimating and testing the logarithm of the rate ratio. Let  $\lambda_0$  and  $\lambda_1$  be the two rate parameters and define

$$\beta_1 = \log(\lambda_1), \quad \beta_0 = \log(\lambda_0)$$

then

$$\begin{aligned} \gamma &= \beta_1 - \beta_0 \\ &= \log\left(\frac{\lambda_1}{\lambda_0}\right) \\ &= \log(\theta), \end{aligned}$$

the log of the rate ratio.

If, in the exposed group,  $D_1$  cases are observed in  $Y_1$  person-years, and in the unexposed group  $D_0$  cases are observed in  $Y_0$  person-years, the total log likelihood is

$$D_1 \log(\lambda_1) - \lambda_1 Y_1 + D_0 \log(\lambda_0) - \lambda_0 Y_0.$$

The gradients of this with respect to  $\beta_1$  and  $\beta_0$  are

$$G_1 = D_1 - \lambda_1 Y_1 \quad G_0 = D_0 - \lambda_0 Y_0,$$

and the curvatures are

$$C_1 = -\lambda_1 Y_1 \quad C_0 = -\lambda_0 Y_0.$$

The most likely values for  $\beta_1$  and  $\beta_0$  are

$$M_1 = \log(D_1/Y_1), \quad M_0 = \log(D_0/Y_0)$$

and the corresponding standard deviations are

$$S_1 = \sqrt{1/D_1}, \quad S_0 = \sqrt{1/D_0}.$$

Using the rules given at the end of the last section, the Gaussian approximation for the profile log likelihood for  $\gamma = \log(\theta)$  has

$$\begin{aligned} M &= \log(D_1/Y_1) - \log(D_0/Y_0) \\ &= \log\left(\frac{D_1/Y_1}{D_0/Y_0}\right), \end{aligned}$$

and

$$S = \sqrt{\frac{1}{D_1} + \frac{1}{D_0}}.$$

These expressions are identical to those obtained in Chapter 13.

The Wald test is also based on the Gaussian approximation shown above. The score test is obtained from the gradient and curvature of the profile log likelihood at the null value of the parameter,  $\gamma = 0$ . Here  $\lambda_1$  and  $\lambda_0$  are equal and their most likely common value is  $D/Y$  so that the gradients and curvatures are

$$\begin{aligned} G_1 &= D_1 - E_1 & G_0 &= D_0 - E_0 \\ C_1 &= -E_1 & C_0 &= -E_0 \end{aligned}$$

where  $E_1 = (D/Y)Y_1$  and  $E_0 = (D/Y)Y_0$  represent 'expected' numbers of failures in the two groups under the null hypothesis. The score,  $U$ , is given by either  $G_1$  or  $-G_0$  (it can easily be verified that these are identical). The score variance is minus the curvature of the profile log likelihood and, using the relationship

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_0}.$$

this is

$$V = \left(\frac{1}{E_1} + \frac{1}{E_0}\right)^{-1}$$

$$= \frac{E_1 E_0}{E}$$

Since  $D = E$ , this can also be written

$$\begin{aligned} V &= D \frac{E_1}{E} \frac{E_0}{E} \\ &= D \frac{E_1}{E} \left(1 - \frac{E_1}{E}\right) \end{aligned}$$

and this agrees with the expression given in Chapter 13.

### THE DIFFERENCE BETWEEN TWO MEANS

A second example is the difference between two mean parameters in a Gaussian model for responses measured on a continuous metric scale. For example, we might wish to compare blood pressure in two groups of subjects. We shall let  $\mu_1$  and  $\mu_0$  represent the mean parameters for the two groups and assume that the standard deviation of responses about the mean is the same in both groups,  $\sigma$  let us say. As in Chapter 8 we shall assume  $\sigma$  to be a known constant although, in practice, it would also have to be estimated from the data.

**Exercise C.1.** Derive expressions for the most likely value and for the standard deviation of the estimate of the parameter

$$\gamma = \mu_1 - \mu_0.$$

### C.2 Weighted sums

Similar results hold for more general problems. For example, the parameter of interest may be defined as

$$\gamma = W_1 \beta_1 + W_0 \beta_0$$

where  $W_1$  and  $W_0$  are known constants. In this case the same argument illustrated in Fig. C.1 may be applied, but the parallel lines corresponding to fixed values of  $\gamma$  now have different slopes. The relationship between gradients in the total log likelihood and the gradient of the profile likelihood is now

$$G = \frac{G_1}{W_1} = \frac{G_0}{W_0}$$

and for the curvatures we have

$$\frac{1}{C} = \frac{(W_1)^2}{C_1} + \frac{(W_0)^2}{C_0}.$$

These results generalize in an obvious way to a function of more than two parameters, of the form

$$\gamma = W_1\beta_1 + W_2\beta_2 + W_3\beta_3 + \dots,$$

the gradient of the profile log likelihood now being

$$G = \frac{G_1}{W_1} = \frac{G_2}{W_2} = \frac{G_3}{W_3} = \dots$$

and its curvature

$$\frac{1}{C} = \frac{(W_1)^2}{C_1} + \frac{(W_2)^2}{C_2} + \frac{(W_3)^2}{C_3} + \dots$$

If the most likely values of  $\beta_1, \beta_2, \dots$  are  $M_1, M_2, \dots$  with standard deviations  $S_1, S_2, \dots$ , then the most likely value of  $\gamma$  is

$$M = W_1M_1 + W_2M_2 + W_3M_3 + \dots$$

with standard deviation

$$S = \sqrt{(W_1S_1)^2 + (W_2S_2)^2 + (W_3S_3)^2 + \dots}$$

### Solutions to the exercises

**C.1** The log likelihoods for  $\mu_1$  and  $\mu_0$  are Gaussian with most likely values  $M_1$  and  $M_0$  — the arithmetic means of the  $N_1$  observations in the first group and the  $N_0$  observations in the second. The corresponding standard deviations are

$$S_1 = \frac{\sigma}{\sqrt{N_1}}, \quad S_0 = \frac{\sigma}{\sqrt{N_0}}.$$

It follows from the results of this section that the profile log likelihood for  $\mu_1 - \mu_0$  has most likely value  $M_1 - M_0$  and standard deviation

$$\sqrt{\frac{(\sigma)^2}{N_1} + \frac{(\sigma)^2}{N_0}} = \sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_0}}.$$

## Appendix D

### Table of the chi-squared distribution

Probability $p$	Degrees of freedom, $\nu$				
	1	2	3	4	5
0.50	0.455	1.386	2.366	3.357	4.351
0.25	1.323	2.773	4.108	5.385	6.626
0.10	2.706	4.605	6.251	7.779	9.2367
0.075	3.170	5.181	6.905	8.496	10.008
0.050	3.841	5.991	7.815	9.488	11.070
0.025	5.024	7.378	9.348	11.143	12.833
0.0100	6.635	9.210	11.345	13.277	15.086
0.0075	7.149	9.786	11.966	13.937	15.780
0.0050	7.879	10.597	12.838	14.860	16.750
0.0025	9.141	11.983	14.320	16.424	18.386
0.0010	10.828	13.816	16.266	18.467	20.515

Probability $p$	Degrees of freedom, $\nu$				
	6	7	8	9	10
0.50	5.348	6.346	7.344	8.343	9.342
0.25	7.841	9.037	10.219	11.389	12.549
0.10	10.645	12.017	13.362	14.684	15.987
0.075	11.466	12.883	14.270	15.631	16.971
0.050	12.592	14.067	15.507	16.919	18.307
0.025	14.449	16.013	17.535	19.023	20.483
0.0100	16.812	18.475	20.090	21.666	23.209
0.0075	17.537	19.229	20.870	22.471	24.038
0.0050	18.548	20.278	21.955	23.589	25.188
0.0025	20.249	22.040	23.774	25.462	27.112
0.0010	22.458	24.322	26.124	27.877	29.588

The above tables give the value that a variable, distributed according to the chi-squared distribution with  $\nu$  degrees of freedom, will exceed with probability  $p$ . For example, a variable distributed according to the chi-squared distribution with one degree of freedom has a probability of  $p = 0.1$  of exceeding the value 2.706.



---

# Index

---

- Aalen-Nelson estimate, 48
- additive model, 224, 282
  - for case-control study, 285
- affected sib pair method, 97
- age-period-cohort model, 315
- age-specific rate, 53
- Akaike's information criterion, 272
- aliasing, 313
- analysis of variance, 339
  
- band (time), 27
- baseline rates, 299, 304
- Bayes' rule, 12
  - and degree of belief, 92
- Bayesian theory of statistics, 22, 92, 117
- binary data, 5
- binary probability model, 6, 18
  - for case-control study, 159
  - sequence of, 27
- binomial distribution, 112, 129
  
- case-cohort study, 162, 331
- case-control study, 153
  - frequency matched, 178
  - geographically based, 163
  - group matched, 178
  - hospital-based, 163
  - individually matched, 178, 186
  - neighbourhood matched, 183
  - size of, 210
- case/control ratio, 158
- cause-specific rate, 65
- cause-specific risk, 63
- censoring, 5, 24, 63
  - non-informative, 68
- chi-squared distribution, 71
- click (time), 42, 146
- cohort study, 5
  - size of, 210
- collinearity, 247
- competing causes, 63
- complementary log-log, 235
- conditional likelihood, 128
- conditional logistic regression, 176, 290
- confidence interval, 21, 90
  - exact, 116, 129, 172
- confounding, 53, 133, 264, 272
- coverage probability, 89, 99
  
- Cox's regression analysis, 298
- credible interval, 22, 94
- cross-validation, 272
- cumulative failure rate, 46, 131
- cumulative survival probability, 29, 46
- curvature, 84, 354, 359
  
- degree of belief, 21, 92
  - posterior, 93
  - prior, 93
- departures from linearity, 252
- deterministic model, 3
- deviance, 242
- differential misclassification, 277
- discriminating between models, 284
- dose-response relationship, 249
  
- effect modification, 276
- endogenous variable, 276
- error factor, 82
- exact confidence interval, 116
- exogenous variable, 276
- expected number of cases, 56, 58, 106, 115, 148, 205
- experiment of nature, 133, 272
- explanation, 271
- exposure, 10, 272
- exposure window, 182
  
- F distribution, 342
- F ratio tests, 342
- factor, 224
- factorial, 115
- failure, 6
- first derivative, 84
- Fisher's exact test, 172
- force of mortality, 40
- frequency record, 225, 227, 346
- frequentist theory of statistics, 21, 89, 173
  
- Gaussian probability model, 71
- Gaussian regression, 336
- genotype, 13
- goodness-of-fit tests, 246
- gradient, 84, 354, 357
  
- haplotype, 13
  - marker, 96

- hazard rate, 40  
 hypergeometric distribution, 170
- incidence, 6  
 incidence density sampling, 161, 182  
 independence  
   of causes, 66  
   of censoring, 68  
   of effects, 282  
   statistical, 11
- indicator variable, 254  
 individual record, 225, 227  
 instantaneous rate, 40  
 interaction, 299, 268, 276, 282  
   between categorical and quantitative, 266  
   between confounders, 261  
   between exposure and confounders, 263  
   scale dependence, 269  
 iterative calculation, 144
- Kaplan-Meier estimate, 35
- late entry, 68, 302  
 least squares estimate, 338  
 Lexis diagram, 57, 228  
 life table, 27  
   actuarial, 27  
   cohort, 31  
   modified, 48  
 likelihood, 18  
   Bernoulli, 19  
   conditional, 128  
   hypergeometric, 171, 176, 188  
   partial, 301, 308, 333  
   Poisson, 44  
   profile, 124  
 likelihood ratio, 20  
 linear effect, 250  
 linkage, genetic, 96  
 lod score  
   and p-value, 107  
 log likelihood, 22, 117  
   approximate, 78  
   for cause-specific rate, 66  
   for rate parameter, 43  
     approximate, 79, 84  
   for risk parameter, 24  
     approximate, 79, 85  
   Gaussian, 74  
 log likelihood ratio, 23, 96  
   test, 100, 237  
 log rank test, 146  
 logistic regression, 176, 202, 227, 229
- Mantel extension test, 203  
 Mantel-Cox test, 146
- Mantel-Haenszel estimate  
   in 1:1 matched studies, 186  
   in 1:m matched studies, 190  
   in case-control studies, 178  
   of rate ratio, 145  
 Mantel-Haenszel test, 177  
   in 1:1 matched studies, 186  
   in 1:m matched studies, 190
- matched sets  
   1:1, 292  
   1:m, 294
- maximum likelihood estimate, 20  
 mean parameter, 73  
   difference between two, 361  
 measurement error  
   confounder, 280  
   exposure, 277  
 missing value, 226  
 Monte Carlo test, 99, 112  
 most likely value, 20  
 multiple regression, 336  
 multiplicative model, 221
- nested case-control study, 162, 329  
   counter-matching in, 332  
   matching in, 332  
   two-stage, 334  
 nested models, 246  
 non-identifiability, 313  
 normal distribution, 71  
 null hypothesis, 96
- observation time, 42, 54  
 Occam's razor, 237  
 odds parameter, 7  
 odds ratio parameter, 155, 161, 166  
   common across strata, 175  
 one-sided test, 105  
 overmatching, 181
- p-value, 99  
   exact, 104, 110, 129, 342  
   misinterpretation of, 107  
   one and two-sided, 105, 112
- parameter, 3  
   corner, 220  
   difference between two, 130, 357  
   estimation, 8, 18  
   interaction, 240  
   location, 73  
   names, 220  
   nuisance, 124  
   null value, 96  
   scale, 73
- person-time, 42  
 Poisson distribution, 115  
 Poisson regression, 198, 227  
 power of study, 206

- prediction, 271  
 prevalence, 6, 235  
   in case-control studies, 164
- probability  
   additive rule, 6  
   conditional, 10, 28  
   marginal, 11, 133  
   multiplicative rule, 11  
   subjective, 92  
 probability rate, 40  
 proband, 25  
 profile log likelihood, 125  
   approximate, 130, 357  
 proportional hazards model, 142, 147
- quadratic curve, 74  
 quadratic dose-response, 253  
 Quetelet's index, 271
- rare disease assumption, 8, 161  
 rate difference parameter, 129, 130, 224  
 rate parameter, 40  
   relationship to risk, 46  
 rate ratio parameter, 124, 161  
   common across strata, 142  
 recall bias, 163  
 reference category, 160  
 reference rates, 58, 106, 147  
 regression model, 217  
 reparametrization, 4, 124  
 residual standard deviation, 337  
 residual sum of squares, 338  
 residual variance, 337  
 risk, 235  
 risk parameter, 7  
   relationship to rate, 46  
 risk ratio parameter, 13, 161  
 risk score, 271  
 risk set, 300
- sampling risk sets, 330  
 saturated model, 242, 339  
 score, 103  
 score test, 100, 102  
 score variance, 103  
 screening  
   predictive value, 13  
   sensitivity, 13  
   sojourn time, 323  
   specificity, 13  
 second derivative, 84  
 selection bias, 162, 183, 309  
   due to censoring, 68  
   due to late entry, 68  
 significance test, 96, 99  
 standard deviation parameter, 73  
 standardization  
   direct, 136  
   indirect, 148  
 standardized mortality ratio, 60, 148  
   comparison of, 149  
 standardized rate, 136  
   comparison of, 139, 149  
 stepwise regression, 271  
 stochastic model, 4  
 stratification, 135  
   in case-control studies, 175, 203  
   in cohort studies, 141, 200  
 study base, 153  
 support, 18, 117  
 supported range, 20  
   approximate, 79  
   for odds parameter  
   approximate, 83  
   for rate parameter, 44  
   approximate, 80, 82  
   for risk parameter, 21  
   approximate, 79, 83  
 survival curve, 32  
 synergism, 282
- time band, 227  
 time scale, 53, 309  
 transformation of parameter, 80, 86  
 trend test, 249  
 trend, testing for, 197, 252
- vague prior, 117  
 variable, 224  
   binary, 225  
   categorical, 224  
   derived, 225  
   explanatory, 219, 272  
   levels, 224  
   quantitative, 224  
   time-varying, 307  
 variable selection strategy, 271  
 variance parameter, 73
- Wald test, 100, 101, 237