random term in the model to estimate the variance associated with each term in the model separately from the other components of variation (Section 9.1.6). The costs (C) must also be determined, preferably from our pilot study where costs can be estimated empirically. The cost for each quadrat is simply the time and/or money required to place the quadrat and estimate the percentage cover of algae, say five minutes. The cost for each patch would be the time taken to move all the gear to each patch (20 minutes) and the time taken to move between quadrats in each patch (three minutes) but NOT the time taken to process a quadrat.

A number of textbooks (Snedecor & Cochran 1989, Sokal & Rohlf 1995, Underwood 1997) provide equations for relating costs and variances to determine the optimum number of replicates at each level of sampling (and see Andrew & Mapstone 1987). In a two factor design, the optimum number of replicates (e.g. quadrats) in each level of B (e.g. each patch) is:

$$n = \sqrt{\frac{C_{B(A)} s^2_{C(B(A))}}{C_{C(B(A))} s^2_{B(A)}}} \tag{9.10}$$

where C is the cost for the appropriate level and s^2 is the estimate of the variance, i.e. the mean square. Note that if the costs of recording a single quadrat are the same as the costs of setting up a new patch, then the sample size is just based on the ratio of the two variance components. Based on the variances and the costs listed above, the optimal number of quadrats per patch is 0.88, i.e. one (Box 9.3).

The number of patches (q) for each density treatment can be determined in two ways based on either the desired variance of the mean for each site (s_A^2) or the fixed total cost of sampling a site (C_A):

$$s_A^2 = \frac{n s^2_{B(A)} + s^2_{C(B(A))}}{nq} \tag{9.11}$$

$$C_A = qC_{B(A)} + nqC_{C(B(A))} \tag{9.12}$$

In the first case, we fix the desired level of precision for the mean of each site (s_A^2) and, using our values for n and the estimated variance components for quadrats and patches, solve for q. In the second case, we fix the total available cost for sampling each density and, again using our values for n and the estimated variance components for quadrats and patches, solve for q. In practice, having a fixed total cost, in time or money, is likely so the latter approach might be used more often. If we set the total cost for setting up each density treatment as four hours (240 minutes), then the number of patches would be 8.6, i.e. nine (Box 9.3). So based on these estimates, the most efficient design would be one quadrat per patch and nine quadrats per treatment. Note that these costs are guesses on our part so we are not suggesting that there was anything wrong with the design used by Andrew & Underwood (1993).

Keough & Mapstone (1995) made a number of sensible recommendations for deriving and using these values for sample size at each level of subsampling. First, the calculated sample sizes depend on the quality of the pilot data, particularly the variance estimates, and how well the variances in the subsequent main study will match those from the pilot study. It is important, therefore, that the pilot study is done in similar locations and at a similar time (e.g. season) to the main study. It is also important to check that these variance estimates still hold once the main research has started and adjust the sample sizes if necessary. It is much easier to reduce sample size during an ongoing research program than to increase them, so the initial sample sizes should be generous. Second, the sample size values will usually not be integers so they should be rounded up to the nearest integer. Finally, the calculations may recommend sample sizes of less than one, because the variance at that level is so small or the costs so cheap. However, some level of replication is necessary for sensible inference and, remembering that pilot studies may underestimate the true variance, we recommend that more than one replicate at any level should always be used.

9.2 Factorial designs

An alternative multifactor linear model is used when our design incorporates two or more factors that are crossed with each other. The term crossed indicates that all combinations of the factors are

Table 9.7 Illustration of marginal and cell means for a two factor factorial ANOVA design. Data from Quinn (1988) where factor A is limpet density, factor B is season and the response variable is number of egg masses per limpet in three replicate enclosures per cell

	B_1	B_2	B_j	Marginal means A
A_1	μ_{11}	μ_{12}	μ_{1j}	$\mu_{i=1}$
A_2	μ_{21}	μ_{22}	μ_{2j}	$\mu_{i=2}$
A_i	μ_{i1}	μ_{i2}	μ_{ij}	μ_i
Marginal means B	$\mu_{j=1}$	$\mu_{j=2}$	μ_j	Grand mean μ

	Factor B (B_j) Season	B_1 Spring	B_2 Summer	Factor A marginal means
Factor A (A_i)	Density			
A_1	8	$\bar{y}_{11} = 2.417$	$\bar{y}_{12} = 1.833$	$\bar{y}_{i=1} = 2.125$
A_2	15	$\bar{y}_{21} = 2.177$	$\bar{y}_{22} = 1.178$	$\bar{y}_{i=2} = 1.677$
A_3	30	$\bar{y}_{31} = 1.565$	$\bar{y}_{32} = 0.811$	$\bar{y}_{i=3} = 1.188$
A_4	45	$\bar{y}_{41} = 1.200$	$\bar{y}_{42} = 0.593$	$\bar{y}_{i=4} = 0.896$
Factor B marginal means		$\bar{y}_{j=1} = 1.840$	$\bar{y}_{j=2} = 1.104$	$\bar{y} = 1.472$

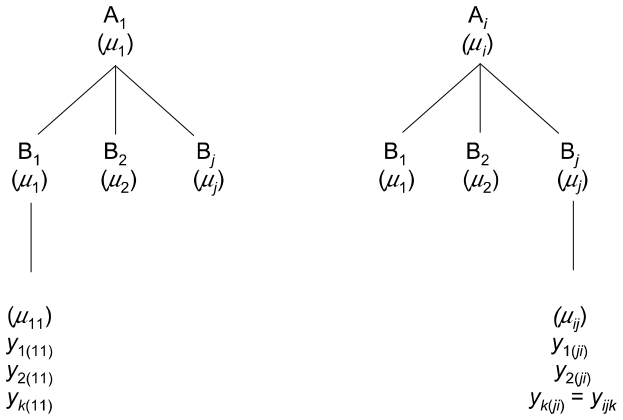

Figure 9.2 Part of data set for two factor crossed ANOVA, with p levels of factor A ($i = 1$ to p), q levels of factor B ($j = 1$ to q), where the levels of B are the same and crossed with each level of A, and n replicate observations within each combination (cell) of A and B ($k = 1$ to n).

included in the design and that every level (group) of each factor occurs in combination with every level of the other factors. Such designs are also termed factorial. This pattern is in contrast to nested designs, where the levels of the nested factor are different within each level of the main factor. We will first consider factorial (crossed) designs with two factors, where every level of one factor occurs at every level of the other factor and both factors are of equal importance – see Figure 9.2 and Table 9.7.

Factorial designs are most often used for manipulative experiments. For example, Poulson & Platt (1996) examined the effects of light microenvironment (three levels: beneath canopy, single treefall gap, multiple treefall gap) and seedling height class (three levels: 1–2 m small, 2–4 m medium, 4–8 m large) on the difference in growth between sugar maple and beech saplings (the response variable was the difference in growth of paired seedlings of each species). There were five replicate seedling pairs for each of the nine microenvironment–height combinations. Another example comes from Maret & Collins (1996), who set up an experiment to test the effects of invertebrate food level and the presence or absence of tadpoles on variation in size among larval salamanders. There were two factors: two levels of ration of invertebrate prey (low and high amounts of brine shrimp per day) and two levels of tadpole supplementation (with and without). There were originally eight replicate aquaria in each of the four cells, although some aquaria were omitted from analysis because one or more salamander larvae died. The response variable was mean snout–vent length of salamanders in each aquarium.

In these two examples, both factors in the design are fixed, i.e. all possible levels of interest for the two factors have been used in the study

and our inference is restricted to these levels. These are analyzed with fixed effects linear models, also termed Model 1 analyses of variance.

Factorial designs can include random factors that are often randomly chosen spatial or temporal units. Designs that include only random factors are analyzed with random effects models, termed Model 2 analyses of variance, although these are unusual in biology. One example is from Kause *et al*. (1999), who examined phenotypic plasticity in the foraging behavior of sawfly larvae with an experiment that used six species of sawflies and 20 individual mountain birch trees that represented a range of leaf qualities for the herbivorous sawfly larvae. There were between four and six larvae per tree and species combination and the response variable was an aspect of foraging behavior (e.g. number of meals, relative consumption rate etc.). Both sawfly species and individual tree were random factors as they were a sample from all possible herbivorous sawflies and all possible trees.

Designs with a combination of fixed and random factors are analyzed with mixed linear models, also termed Model 3 analyses of variance. Including a random factor in a multifactor design is important in biology, because it allows us to generalize the effects of a fixed factor to the population of spatial or temporal units (Beck 1997). For example, Brunet (1996) tested the effects of position on an inflorescence and randomly chosen plants on fruit and seed production of a perennial herb. This was a two factor design with flower position as the fixed factor and individual plant as the random factor. A second example comes from Twombly (1996), who randomly assigned copepod nauplii from 15 sibships to one of four food treatments (high constant food and high switched to low at three different naupliar stages); there were four replicate dishes (each containing two nauplii) per factor combination and the response variable was age at metamorphosis. Food treatment was a fixed factor and sibship was a random factor.

Factorial designs can include three or more factors (Section 9.2.12), although we will illustrate the principles based on two factor designs. Factorial designs allow us to measure two different sorts of factor effects.

1. The main effect of each factor is the effect of each factor independent of (pooling over) the other factors.

2. The interaction between factors is a measure of how the effects of one factor depend on the level of one or more additional factors. The absence of an interaction means that the combined effect of two or more factors is predictable by just adding their individual effects together. The presence of an interaction indicates a synergistic or antagonistic effect of the two factors.

We can only measure interaction effects in factorial (crossed) designs. In nested designs where factor B is nested within factor A, different levels of B are used in each level of A so any interaction between A and B cannot be assessed. When all possible combinations of the two (or more) factors are used in factorial designs they are called complete factorials. Sometimes this is logistically impossible because the experiment would be too big and/or costly, so a subset of factor combinations is used and the design is termed a fractional factorial. Such designs are more difficult to analyze because not all interactions can be measured – see Section 9.2.12.

Fecundity of limpets: effects of season and adult density

Our first worked example of a factorial ANOVA design and analysis is from Quinn (1988). He examined the effects of season (two levels, winter/spring and summer/autumn) and adult density (four levels, 8, 15, 30 and 45 animals per 225 cm^2) on the production of egg masses by rocky intertidal pulmonate limpets (*Siphonaria diemenensis*). Limpets (approx. 10 mm shell length) were enclosed in 225 cm^2 stainless steel mesh enclosures attached to the rocky platform. There were eight treatment combinations (four densities at each of two seasons) and three replicate enclosures per treatment combination. Note that all four densities were used in both seasons, hence a factorial or crossed design. One of the important questions being asked with this experiment was whether the effect of density on number of egg masses per limpet depended on season. Quinn (1988) predicted that the density effect would be greater in summer/autumn, when algal food was

scarce, than in winter/spring, when algal food was more abundant.

Quinn (1988) described another experiment looking at the same species of limpet lower on the shore. Here the limpets were bigger (15–20 mm shell length) and there was much less seasonal variation in the availability of algal food, algal cover being high all year round. The same two factors were used for this experiment but only three densities were included: 6, 12 and 24 limpets per 225 cm^2. So there were six treatment combinations (three densities at each of two seasons) and three replicate enclosures per treatment combination. The analyses of both experiments are in Box 9.4.

Oysters, limpets and mangrove forests

Our second example is from Minchinton & Ross (1999), who examined the distribution of oysters, and their suitability as habitat for limpets in a

Box 9.4 Worked example of two factor fixed effects ANOVA

Quinn (1988) examined the effects of season (winter/spring and summer/autumn) and adult density (8, 15, 30 and 45 animals per 225 cm^2 enclosure) on the production of egg masses by intertidal pulmonate limpets (*Siphonaria diemenensis*). There were three replicate enclosures per treatment combination and the response variable was the number of egg masses per limpet in each enclosure.

The null hypotheses were as follows.

No difference between mean number of egg masses laid in each season, pooling densities.

No difference in mean number of egg masses laid at each density, pooling seasons.

No interaction between season and density, i.e. the effect of density on mean numbers of egg masses laid is independent of season and vice versa.

Source	df	MS	*F*	*P*
Density	3	1.76	9.67	0.001
Linear	1	5.02	27.58	<0.001
Quadratic	1	0.24	1.29	0.272
Season	1	3.25	17.84	0.001
Density × season	3	0.06	0.30	0.824
Residual	16	0.18		

There were no outliers and the residual plot (Figure 9.4(a)) did not suggest problems with assumptions. There was no evidence of an interaction ($P = 0.824$, see Figure 9.5(a)). There were significant effects of season (more egg masses in winter/spring than summer/autumn) and density. The main effect of density was further analyzed with orthogonal polynomials (see Chapter 8 and Section 9.2.10). There was a significant negative linear trend in egg mass production with density but no quadratic trend.

Quinn (1988) did a similar experiment at a lower level of the same shore where the limpets were larger. Different densities were used (6, 12, 24) but the same two seasons with three replicate enclosures per treatment combination. The null hypotheses were the same as above, except that there were only three densities. Again, the residual plot did not suggest any problem with variance heterogeneity (Figure 9.4(b)).

Source	df	MS	*F*	*P*
Density	2	2.00	13.98	0.001
Season	1	17.15	119.85	<0.001
Density × season	2	0.85	5.91	0.016
Density 6 vs 12 & 24 × season	1	1.53	10.66	0.007
Linear density × season	1	1.44	10.07	0.008
Residual	12	0.14		

There was a significant interaction between density and season ($P = 0.016$, Figure 9.5(b)). Treatment–contrast interaction tests showed that the comparison between control density and increased density varied between seasons and the linear trend in density was also significantly different between seasons. We also tested simple main effects of density separately for each season.

Source	df	MS	*F*	*P*
Winter density	2	0.17	1.21	0.331
Summer density	2	2.67	18.69	<0.001
Residual	12	0.14		

The effect of density was only significant in summer, not in winter. Note that the original $MS_{Residual}$ was used for both tests.

temperate mangrove forest. They chose two sites about 600 m apart and at each site recorded the density of oysters in four zones running up the shore: seaward zone without mangrove trees, seaward zone with mangrove trees, middle zone with trees, and a landward zone at the upper levels. In each of the eight combinations of site and zone, they used five quadrats to sample oysters (response variable) on the forest floor. An additional study examined the distribution of limpets on oysters on bent mangrove tree trunks. They used two sites, three zones (obviously the seaward zone without trees was not included) and two orientations of mangrove trunk (upper facing canopy and lower facing forest floor). This was a three factor sampling design with five quadrats in each of the 12 cells and densities of limpets per oyster surface as the response variable. For both designs, site was a random factor, representing all possible sites within the mangrove forest, and zone and orientation were fixed factors. The analyses of these data are in Box 9.5.

9.2.1 Linear models for factorial designs

In the sections that follow, we will describe two factor designs and their associated linear models. Designs with more than two factors will be examined in Section 9.2.12. A two factor factorial design is illustrated in Figure 9.2 with a factor relationship diagram. Factor A has p groups ($i = 1$ to p), factor B has q groups ($j = 1$ to q) crossed with each level of A and there are n_i replicates ($k = 1$ to n_i) within each combination of A and B categories, i.e. each cell. Note that every level of factor B is crossed with every level of factor A and vice versa. For the moment, assume the number of replicate observations (n) in each combination of A and B is the same. Unequal sample sizes will be discussed in Section 9.2.6. There will be a total of pq cells in this factorial design with n replicate observations in each cell. From Quinn (1988), p was four limpet density treatments (factor A), q was two seasons (factor B) and n was three enclosures within each cell. From Minchinton & Ross (1999), p was four zones (factor A), q was two sites (factor B) and n was five quadrats within each cell.

We need to distinguish between two types of means in multifactor crossed designs (Table 9.7).

- Marginal means are the means for the levels of one factor pooling over the levels of the