

For questions 1–3, please also refer to sample analyses of the data attached at the back of these solutions.

1. Draw the Hasse diagram for this experiment.

See below.

2. What are the appropriate test denominators for the parental lines, the inbred tester lines, and their interaction?

The denominators are location by parental lines, location by inbred tester, and location by parental by inbred, respectively.

3. Describe your conclusions for this experiment with respect to the parental lines, the inbred tester lines, and their interaction.

There are no main effects, but some individual combinations differ from the others. In particular, two combinations involving the second inbred tester are unusually low in response.

(4) A sociologist is developing a new questionnaire ...

The two questionnaires are the treatments. We need to block on major and age, so we could use a two by two Latin Square replicated 20 times.

5) Air flow through heating and air conditioning vents can become noisy...

No indication of how we might block here, so I would do a completely randomized design with vent as unit and vent design as treatment.

(6) One of the problems encountered when restoring a wetland ...

This is a randomized complete block. The sites are the blocks, and the plots are the units. There are eight treatments in 2^3 structure.

Source	DF
Block	1
burn	1
till	1
herb	1
burn.till	1
burn.herb	1
till.herb	1
burn.till.herb	1
Error	7

7) A consumer testing agency is trying to compare four over-the-counter ...

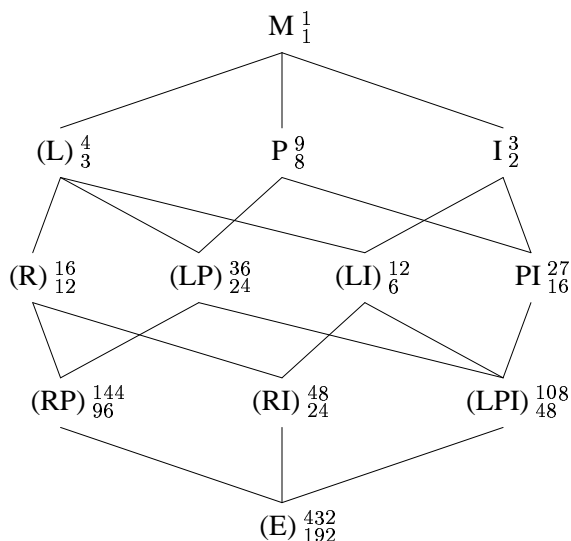
This is a balanced incomplete block. The subjects are the blocks, and the face halves are the units.

The creams are the treatments.

Source	DF
Block	95
cream	3
Error	93

Summary. No single parental line or inbred tester line produces consistently higher or lower responding hybrids. However, some of the combinations using inbred tester two are unusually low. In addition, there is considerable evidence that the parental lines, inbred testers, and hybrids have different responses at different locations.

Analysis. This experiment is a randomized complete block. The replications at each of the locations (16 replications total) are the blocks, and the 27 parental line by inbred tester line combinations (in 9 by 3 factorial structure) are the treatments. The blocks may be split into locations and replications nested in locations. The locations are random, and presumably the replications are as well. Prior belief indicates that the treatments may interact with the locations. Replications might interact with treatments, but we are not given guidance in that matter. Fortunately, inclusion or exclusion of replication by treatment interactions will not affect the tests for treatments (though it will affect the tests for main effect by location interaction variance components). Here we include the replication by treatment interactions. The Hasse diagram follows:



The expected mean squares are:

$$\begin{aligned}
 \text{EMS}(L) &= V(E) + 27V(R) + 108V(L) \\
 \text{EMS}(R) &= V(E) + 27V(R) \\
 \text{EMS}(P) &= V(E) + 3V(R.P) + 12V(L.P) + 48Q(P) \\
 \text{EMS}(L.P) &= V(E) + 3V(R.P) + 12V(L.P) \\
 \text{EMS}(R.P) &= V(E) + 3V(R.P) \\
 \text{EMS}(I) &= V(E) + 9V(R.I) + 36V(L.I) + 144Q(I) \\
 \text{EMS}(L.I) &= V(E) + 9V(R.I) + 36V(L.I) \\
 \text{EMS}(R.I) &= V(E) + 9V(R.I) \\
 \text{EMS}(P.I) &= V(E) + 4V(L.P.I) + 16Q(P.I) \\
 \text{EMS}(L.P.I) &= V(E) + 4V(L.P.I) \\
 \text{EMS}(E) &= V(E)
 \end{aligned}$$

Residual analysis does not indicate any problems. The mixed effects ANOVA is

	DF	MS	Error DF	Error MS	F	P value
L	3	8.449e+04	12	1.191e+04	7.095	0.00535
R	12	1.191e+04	192	822.3	14.48	0
P	8	1.026e+04	24	6146	1.669	0.158
L.P	24	6146	96	1279	4.804	1.566e-08
R.P	96	1279	192	822.3	1.556	0.005108
I	2	3.824e+04	6	9259	4.13	0.07449
L.I	6	9259	24	1227	7.549	0.0001257
R.I	24	1227	192	822.3	1.492	0.07363
P.I	16	9143	48	1354	6.75	1.184e-07
L.P.I	48	1354	192	822.3	1.647	0.009839
ERROR1	192	822.3	0	0	MISSING	MISSING

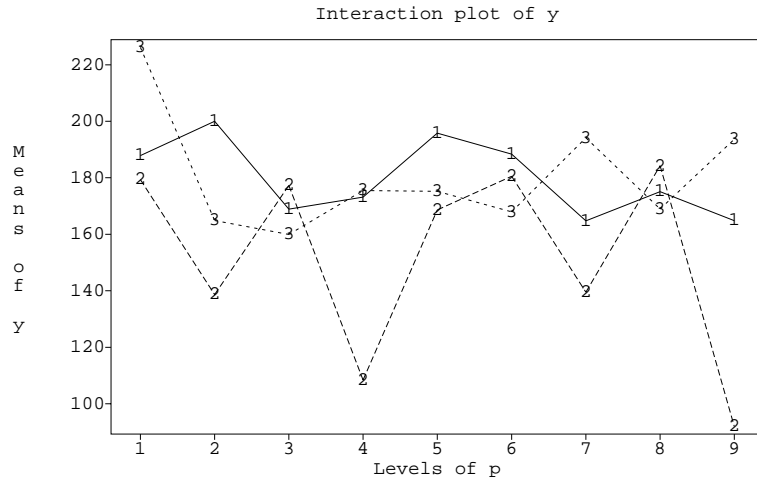
The ANOVA estimates of the variance components are

	Estimate	SE	DF
L	672.07	640.36	2.203
R	410.61	180.09	10.397
L.P	405.57	148.65	14.887
R.P	152.36	67.615	10.155
L.I	223.13	148.82	4.4959
R.I	44.914	40.432	2.468
L.P.I	133.04	72.235	6.7841
ERROR1	822.33	83.929	192

There are location and replication effects, as expected; blocking was successful. There are also statistically significant location by treatment interactions, indicating that the effects of the treatment combinations will vary from location to location. Though (highly) significant, the estimated location by treatment variance components are smaller than the error mean square. For example, a random location will affect the response of a parental line by about plus or minus 20 ($\approx \sqrt{405}$) and affect the response of an inbred line by about plus or minus 15; the experimental error is about plus or minus 29. The (perhaps questionable) interactions involving the replications are smaller still and of mixed significance.

Neither the main effect of parental line nor the main effect of inbred tester line is significant, but the parental line by inbred tester line interaction is highly significant (p-value about .0000001). In large part the lack of significance for main effects is because they have considerably larger denominators in their F-ratios (LP or LI instead of LPI); that is, there are some large main effects but they are not significant when compared with their large standard errors. With a significant interaction and apparently no significant main effects, we see that no parental line gives consistently high or low yields across all tester lines, and no inbred tester line gives consistently high or low yields across all parental lines, but some parental/inbred combinations are high yielders and some are low yielders.

To compare the hybrids consider the interaction plot:



Using methods in Section 12.7, we find that the variance of a treatment mean is

$$\sigma_L^2/4 + \sigma_R^2/16 + \sigma_{LP}^2/4 + \sigma_{LI}^2/4 + \sigma_{RP}^2/16 + \sigma_{RI}^2/16 + \sigma_{LPI}^2/4 + \sigma^2/16.$$

The covariance of two means with the same level of P but different levels of I is

$$\sigma_L^2/4 + \sigma_R^2/16 + \sigma_{LP}^2/4 + \sigma_{RP}^2/16 + \sigma^2/16.$$

The variance of a difference of two such means is

$$\begin{aligned} V(\bar{y}_{ij} - \bar{y}_{ik}) &= 2(\sigma_{LI}^2/4 + \sigma_{RI}^2/16 + \sigma_{LPI}^2/4 + \sigma^2/16) \\ &= \frac{1}{8}(4\sigma_{LI}^2 + \sigma_{RI}^2 + 4\sigma_{LPI}^2 + \sigma^2) \\ &= \frac{1}{72}(EMS_{LI} + 9EMS_{LPI} - EMS_E) \end{aligned}$$

We estimate this variance as

$$\frac{1}{72}(MS_{LI} + 9MS_{LPI} - MS_E)$$

and compute approximate degrees of freedom using the Satterthwaite approximation. The estimated variance is 286.5 with 24.5 degrees of freedom (standard error 16.9).

The covariance of two means with the same level of I but different levels of P is

$$\sigma_L^2/4 + \sigma_R^2/16 + \sigma_{LI}^2/4 + \sigma_{RI}^2/16.$$

The variance of a difference of two such means is

$$\begin{aligned} V(\bar{y}_{ij} - \bar{y}_{kj}) &= 2(\sigma_{LP}^2/4 + \sigma_{RP}^2/16 + \sigma_{LPI}^2/4 + \sigma^2/16) \\ &= \frac{1}{8}(4\sigma_{LP}^2 + \sigma_{RP}^2 + 4\sigma_{LPI}^2 + \sigma^2) \\ &= \frac{1}{24}(EMS_{LP} + 3EMS_{LPI} - EMS_E) \end{aligned}$$

We estimate this variance as

$$\frac{1}{24}(MS_{LP} + 3MS_{LPI} - MS_E)$$

and compute approximate degrees of freedom using the Satterthwaite approximation. The estimated variance is 391.14 with 45.9 degrees of freedom (standard error 19.8).

The covariance of two means with different levels of P and I is

$$\sigma_L^2/4 + \sigma_R^2/16.$$

The variance of a difference of two such means is

$$\begin{aligned} V(\bar{y}_{ij} - \bar{y}_{kl}) &= 2(\sigma_{LP}^2/4 + \sigma_{LI}^2/4 + \sigma_{RP}^2/16 + \sigma_{RI}^2/16 + \sigma_{LPI}^2/4 + \sigma^2/16) \\ &= \frac{1}{8}(4\sigma_{LP}^2 + 4\sigma_{LI}^2 + \sigma_{RP}^2 + \sigma_{RI}^2 + 4\sigma_{LPI}^2 + \sigma^2) \\ &= \frac{1}{72}(3EMS_{LP} + EMS_{LI} + 9EMS_{LPI} - 4EMSE) \end{aligned}$$

We estimate this variance as

$$\frac{1}{72}(3MS_{LP} + MS_{LI} + 9MS_{LPI} - 4MSE)$$

and compute approximate degrees of freedom using the Satterthwaite approximation. The estimated variance is 508.32 with 42.5 degrees of freedom (standard error 22.5).

Returning to the interaction plot, differences between means in the same column (same P different I) have a standard error of 16.9, differences of means on the same line but in different columns (same I different P) have a standard error of 19.8, and all other differences have a standard error of 22.5. The treatment means are

187.9	200.0	168.9	173.2	195.8	188.4	164.8	175.1	164.8
179.4	138.6	177.3	108.4	168.6	180.6	139.5	184.3	91.9
226.2	164.9	159.9	175.5	175.2	167.9	194.2	168.9	193.6

Using a Bonferroni correction for all pairwise comparisons and the smallest of the three equivalent degrees of freedom, means need to be about 4.5 standard errors apart to be significantly different (4.1 for the greatest degrees of freedom). For differences in a column, this is 76; for differences in a row, this is 89; for other differences, this is 101. There are several differences. The value in row 2 column nine is significantly different from the 193 in its column, the 184 in its row, and the means of 226, 200, 195.8, and 194. The high value of 226 is significantly different from only the 108.6 and 91.9 values in the second row. No other differences are significant. Mostly what we see here is two very low values in the second row, and a couple of moderately but not extremely high values.

The differences suggest that the second row may be unusual. A contrast between levels of I with coefficients 1, -2, 1 captures 99.93% of the variation between rows. This would be significant by itself, but is not significant after accounting for data snooping.