

MATH 204: PRINCIPLES OF STATISTICS 2

WINTER 2026

PARAMETER ESTIMATION FOR THE TWO-WAY LAYOUT

For the two-way layout for factors A and B, we may use the data generating model

$$X_{ijt} \sim \text{Normal}(m_{ij}, \sigma^2)$$

for

- $i = 1, \dots, a$ levels of Factor A
- $j = 1, \dots, b$ levels of Factor B
- $t = 1, \dots, r$ for the balanced replication.

with the mean model for m_{ij} broken down into components as follows

$$\begin{aligned} m_{ij} &= \beta_1 & i = 1, j = 1 \\ &= \beta_1 + \beta_i^{(A)} & i = 2, \dots, a, j = 1 \\ &= \beta_1 + \beta_j^{(B)} & i = 1, j = 2, \dots, b \\ &= \beta_1 + \beta_i^{(A)} + \beta_j^{(B)} + \gamma_{ij}^{(AB)} & i = 2, \dots, a, j = 2, \dots, b \end{aligned}$$

The data fall into cells of the $a \times b$ table, where the cell means are m_{ij} . For example, for $a = 3, b = 4$.

		Factor B			
		1	2	3	4
Factor A	1	m_{11}	m_{12}	m_{13}	m_{14}
	2	m_{21}	m_{22}	m_{23}	m_{24}
	3	m_{31}	m_{32}	m_{33}	m_{34}

that is

		Factor B			
		1	2	3	4
Factor A	1	β_1	$\beta_1 + \beta_2^{(B)}$	$\beta_1 + \beta_3^{(B)}$	$\beta_1 + \beta_4^{(B)}$
	2	$\beta_1 + \beta_2^{(A)}$	$\beta_1 + \beta_2^{(A)} + \beta_2^{(B)} + \gamma_{22}^{(AB)}$	$\beta_1 + \beta_2^{(A)} + \beta_3^{(B)} + \gamma_{23}^{(AB)}$	$\beta_1 + \beta_2^{(A)} + \beta_4^{(B)} + \gamma_{24}^{(AB)}$
	3	$\beta_1 + \beta_3^{(A)}$	$\beta_1 + \beta_3^{(A)} + \beta_2^{(B)} + \gamma_{32}^{(AB)}$	$\beta_1 + \beta_3^{(A)} + \beta_3^{(B)} + \gamma_{33}^{(AB)}$	$\beta_1 + \beta_3^{(A)} + \beta_4^{(B)} + \gamma_{34}^{(AB)}$

This model has ab parameters.

Example: Suppose we use the following data generating parameter values

$$\begin{aligned}\beta_1 &= 10 \\ (\beta_2^{(A)}, \beta_3^{(A)}) &= (1, 2.5) \\ (\beta_2^{(B)}, \beta_3^{(B)}, \beta_4^{(B)}) &= (2, 3, 4, 5) \\ (\gamma_{23}^{(AB)}, \gamma_{32}^{(AB)}, \gamma_{34}^{(AB)}) &= (2, -3, 1)\end{aligned}$$

with all other $\gamma_{ij}^{(AB)} = 0$.

```
> set.seed(230)
> a<-3
> b<-4
> r<-6
> n<-a*b*r
> A<-rep(rep(1:a, each=b), r)
> B<-rep(rep(1:b, a), r)
> rept<-rep(1:r, b*a)
> beta1<-10
> betaA<-c(0, 1, 2.5)
> betaB<-c(0, 2:4)
> gammaAB<-matrix(0, nrow=a, ncol=b)
> gammaAB[2, 3]<-2
> gammaAB[3, 2]<- -3
> gammaAB[3, 4]<-1
>
> muvec<-beta1+betaA[A]+betaB[B]
> for(i in 1:n){
+   muvec[i]<-muvec[i]+gammaAB[A[i], B[i]]
+ }
> sig<-2
> X<-muvec+rnorm(n)*sig
> mumat<-tapply(muvec, INDEX=list(A, B), mean)
> mumat
```

	1	2	3	4
1	10.0	12.0	13.0	14.0
2	11.0	13.0	16.0	15.0
3	12.5	11.5	15.5	17.5

The mean structure is

		Factor B			
Factor A		1	2	3	4
	1	10	12	13	14
	2	11	13	16	15
	3	12.5	11.5	15.5	17.5

and there are $r = 6$ replicates. We take $\sigma^2 = 4$.

```
> xbarmat<-tapply(X,INDEX=list(A,B), mean)
> xbarmat

+           1           2           3           4
+ 1   9.68457 12.0792 13.6166 14.7469
+ 2 11.14006 12.8555 16.7459 14.5653
+ 3 12.24099 10.7025 15.1024 16.8112
```

The observed data means are

		Factor B			
		1	2	3	4
Factor A	1	9.68	12.08	13.62	14.75
	2	11.14	12.86	16.75	14.57
	3	12.24	10.70	15.10	16.81

For the ANOVA table we have:

```
> FDdat<-data.frame(A=as.factor(A),B=as.factor(B),X)
> print(summary(aov(X~A+B+A:B, data=FDdat)),digits=5)

+           Df   Sum Sq Mean Sq F value    Pr(>F)
+ A             2    24.702   12.351   3.6885  0.03085 *
+ B             3   268.920   89.640  26.7702 4.094e-11 ***
+ A:B           6    57.377    9.563   2.8558  0.01633 *
+ Residuals    60   200.910    3.349
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

For the parameter estimates:

```
> summary(lm(X~A+B+A:B, data=FDdat))$coef

+           Estimate Std. Error   t value    Pr(>|t|)
+ (Intercept)  9.684569    0.74705 12.963750 4.76445e-19
+ A2           1.455488    1.05649  1.377666 1.73423e-01
+ A3           2.556421    1.05649  2.419735 1.85786e-02
+ B2           2.394591    1.05649  2.266558 2.70357e-02
+ B3           3.932000    1.05649  3.721764 4.38371e-04
+ B4           5.062327    1.05649  4.791655 1.12655e-05
+ A2:B2       -0.679101    1.49410 -0.454522 6.51093e-01
+ A3:B2       -3.933073    1.49410 -2.632403 1.07642e-02
+ A2:B3        1.673835    1.49410  1.120296 2.67052e-01
+ A3:B3       -1.070594    1.49410 -0.716548 4.76433e-01
+ A2:B4       -1.637037    1.49410 -1.095668 2.77603e-01
+ A3:B4       -0.492072    1.49410 -0.329343 7.43043e-01
```

For 95% confidence intervals:

```
> confint(lm(X~A+B+A:B, data=FDdat))
```

```
+           2.5 %      97.5 %
+ (Intercept)  8.190247 11.178892
+ A2          -0.657803  3.568779
+ A3           0.443130  4.669712
+ B2           0.281300  4.507883
+ B3           1.818709  6.045291
+ B4           2.949036  7.175618
+ A2:B2       -3.667746  2.309544
+ A3:B2       -6.921718 -0.944428
+ A2:B3       -1.314810  4.662480
+ A3:B3       -4.059239  1.918051
+ A2:B4       -4.625682  1.351608
+ A3:B4       -3.480717  2.496573
```

Parameter	True value	Estimate	Confidence Interval
β_1	10	9.68	(8.19, 11.18)
$\beta_2^{(A)}$	1	1.46	(-0.66, 3.57)
$\beta_3^{(A)}$	2.5	2.56	(0.44, 4.67)
$\beta_2^{(B)}$	2	2.39	(0.28, 4.51)
$\beta_3^{(B)}$	3	3.93	(1.82, 6.05)
$\beta_4^{(B)}$	4	5.06	(2.95, 7.18)
$\gamma_{22}^{(AB)}$	0	-0.68	(-3.67, 2.31)
$\gamma_{32}^{(AB)}$	-3	-3.93	(-6.92, -0.94)
$\gamma_{23}^{(AB)}$	2	1.67	(-1.31, 4.66)
$\gamma_{33}^{(AB)}$	0	-1.07	(-4.06, 1.92)
$\gamma_{24}^{(AB)}$	0	-1.64	(-4.63, 1.35)
$\gamma_{34}^{(AB)}$	1	-0.49	(-3.48, 2.50)

Finally, we may estimate σ^2 by the MSE from the ANOVA table

$$\hat{\sigma}^2 = 3.349.$$