

# MODELOS EMPÍRICOS

Profa. Dra. Kelly Johana Dussán Medina

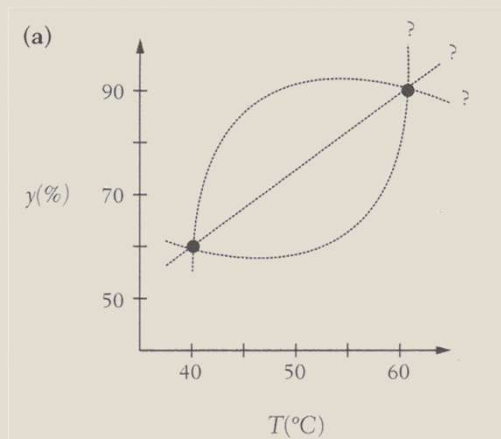
kelly.medina@unesp.br



1

## ¿Cómo construir Modelos Empíricos?

- Consideremos la variación del rendimiento de la reacción con la temperatura. Los rendimientos medios observados con el catalizador A son del 59% y del 90% a 40°C y 60°C, respectivamente.
- Colocando estos dos valores en un gráfico, tienen funciones infinitas.
- Si tomamos 3 mediciones más y vemos que el gráfico de 5 puntos se parece a una línea recta, entonces tenemos más confianza en el modelo.
- Diseños experimentales de 2 niveles es sólo un paso inicial.



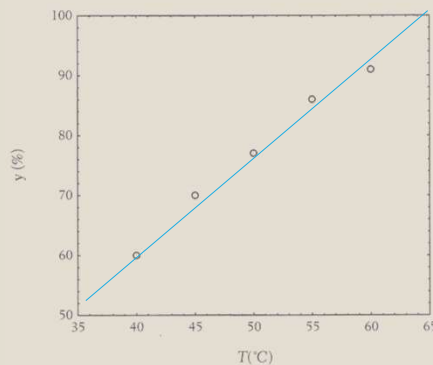
2

2

## Un modelo para $y = f(T)$

### Ejemplo 01

Temperatura (°C)	40	45	50	55	60
Rendimiento (%)	60	70	77	86	91



$$y_i = b_0 + b_1 T_i + \varepsilon_i$$

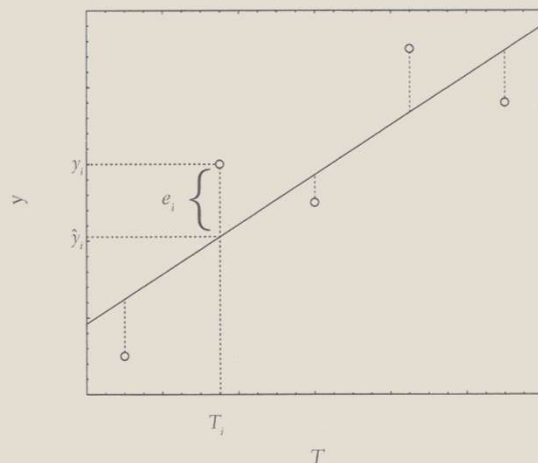


Error aleatorio

3

3

- ¿Cuál es la mejor recta?
- Cualquier recta que decidamos elegir dejará residuos en relación con algunas observaciones.
- Los residuos pueden ser positivos o negativos, dependiendo de si los rendimientos observados están por encima o por debajo de la recta elegida.
- La mejor línea será la que pase más cerca de los puntos experimentales.
- Lo ideal es minimizar la distancia total de los puntos a la recta.



4

4

- La forma tradicional de conseguirlo es situar la línea de forma que la suma de los cuadrados de los residuos sea mínima, por lo que este método se denomina **ajuste de mínimos cuadrados**. También se conoce como **análisis de regresión**.
- Si, a la temperatura  $T_i$ , el rendimiento observado es  $y_i$  y el rendimiento predicho por la recta de regresión es  $\hat{y}_i$ , el residuo que deja el modelo es

$$\varepsilon_i = y_i - \hat{y}_i$$

↑
← Previsto  
Observado

5

5

## Análisis de varianza

- El estudio de los residuos es fundamental para evaluar la calidad del ajuste de cualquier modelo.
- Los residuos deben ser pequeños.
- En el modelo ideal, todas las predicciones coincidirían exactamente con las respuestas observadas y no habría ningún residuo.
- El método más utilizado para evaluar numéricamente la calidad del ajuste de un modelo es el **análisis de la varianza**.

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

↑
← Error Aleatorio  
Desviación por modelo

6

6

- El siguiente paso consiste en expresar esta comparación de desviaciones en términos cuantitativos. Para ello, elevamos la ecuación al cuadrado y sumamos todos los puntos.

$$\sum (y_i - \bar{y})^2 = \sum [(\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)]^2$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

- Estas sumas de cuadrados de desviaciones suelen denominarse sumas de cuadrados o SQ

$$[\text{SQ en torno a la media}] = [\text{SQ por regresión}] + [\text{SQ residual}]$$

$$SQ_T = SQ_R + SQ_r$$

7

7

- Parte de la variación total de las observaciones  $y_i$  en torno a la media  $\bar{y}$  es descrita por la ecuación de regresión, y el resto se explica por los residuos.
- Cuanto mayor sea la fracción descrita por la regresión, mejor será el ajuste del modelo.

$$R^2 = \frac{SQ_R}{SQ_T} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

- $R^2$  se denomina coeficiente de determinación del modelo. El valor máximo de  $R^2$  es 1 y sólo se producirá en el improbable caso de que no haya ningún residuo.
- Cuanto más se acerque el valor  $R^2_{\text{adj}}$  a 1, mejor se ajustará el modelo a las respuestas observadas.

$$R^2_{\text{ajustado}} = 1 - \frac{n-1}{n-p} (1 - R^2)$$

8

8

- $R^2$  representa el porcentaje de variación de la respuesta que explica el modelo. Cuanto mayor sea el valor  $R^2$  mejor se ajusta el modelo a los datos.  $R^2$  está siempre entre 0 y 100%.

$$R^2 = \frac{SQ_R}{SQ_T}$$

- $R^2_{ajustado}$ , determina el grado de varianza de la variable dependiente que puede explicar la variable independiente. Observando el valor  $R^2_{ajustado}$ , puede evaluarse si los datos de la ecuación de regresión están correctamente ajustados. Cuanto mayor sea la  $R^2_{ajustado}$ , mejor será la ecuación de regresión, ya que implica que la variable independiente elegida para determinar la variable dependiente puede explicar la variación de la variable dependiente.

$$R^2_{adj} = 1 - \left[ \frac{(n-1)}{(n-p)} \times (1 - R^2) \right]$$

- % máxima variación explicable

$$\%mve = 1 - \frac{SQ_{ep}}{SQ_T}$$

9

9

- A cada suma cuadrática se le asocia un cierto número de grados de libertad, que indica cuántos valores independientes de las  $n$  observaciones  $y_n$  se necesitan para determinarla.
- Para la suma cuadrática de las "n" desviaciones de la media, el número de grados de libertad es  $(n-1)$ . Esto se debe a que la suma de las desviaciones es cero y esto consume un grado de libertad.
- La suma cuadrática debida a la regresión sólo tiene un grado de libertad. Como el número de grados de libertad de la  $SQ_T$  es  $(n-1)$ , la suma cuadrática residual debe tener  $(n-2)$  grados de libertad para satisfacer la siguiente ecuación:

$$v_T = v_R + v_r$$

$$(n-1) = 1 + (n-2)$$

El lado derecho refleja el hecho de que nuestro modelo sólo contiene dos parámetros.

10

10

En el caso general de un modelo con "p" parámetros, el número de grados de libertad de la suma cuadrática residual viene dado por la diferencia entre el número de observaciones y el número de parámetros estimados

$$v_r = (n - p)$$

$$v_T = (n - 1)$$

$$v_R = (p - 1)$$

r = residual; R = Regresión, T = total

n = número de experimentos; p = número de parámetros del modelo.

11

11

- Estos resultados pueden representarse en la denominada tabla de análisis de la varianza o, simplemente, **ANOVA**.
- Dividiendo las sumas cuadráticas por sus respectivos números de grados de libertad, obtenemos las llamadas **medias cuadráticas** (MQ).

Fuente de Variación	Suma Cuadrática	G.L.	Media Cuadrática
Regresión	$\sum (\hat{y}_i - \bar{y})^2$	p-1	$MQ_R = SQ_R / (p-1)$
Residuos	$\sum (y_i - \hat{y}_i)^2$	n-p	$MQ_r = SQ_r / (n-p)$
Total	$\sum (y_i - \bar{y})^2$	n-1	

n = ensayos y p = parámetros  
 $\hat{y}_i$  Previsto;  $y_i$  Observado;  $\bar{y}$  media.

12

12

Modelo:

$$y = -1,2 + 1,56 T$$

$$\text{Media: } \bar{y} : 76,80$$

Temperatura (°C)	40	45	50	55	60
Rendimiento Exp., $y_i$	60	70	77	86	91
Rendimiento Modelo, $\hat{y}_i$	61,2	69,0	76,8	84,6	92,4

$$SQ_R = \sum (\hat{y}_i - \bar{y})^2 \quad SQ_r = \sum (y_i - \hat{y}_i)^2 \quad Total = \sum (y_i - \bar{y})^2$$

Ensayos	$SQ_R$	$SQ_r$	Total	Fuente de Variación	Suma Cuadrática	G.L.	Media Cuadrática
1	243,36	1,44	282,24	Regresión	608,4	1	608,4
2	60,84	1,00	46,24	Residuos	6,4	3	2,13
3	0	0,04	0,04	Total	614,8	4	
4	60,84	1,96	84,64				
5	243,36	1,96	201,64				
Total	608,4	6,40	614,8				

$p = 2; n = 5$

13

13

- $MQ_R$  es una estimación de la varianza de los puntos en torno al modelo..
- $MQ_r$  puede interpretarse como una medida del error cuadrático medio de la ecuación de regresión.
- Las medias cuadráticas se utilizan para evaluar si la ecuación de regresión es estadísticamente significativa.
- Para la prueba de significancia se utiliza una distribución F.

$$F_{GL_R, GL_r} = \frac{MQ_R}{MQ_r}$$

$$F_{\text{calculado}} > F_{\text{tabulado}}$$



**El modelo de regresión es estadísticamente significativo**

14

14

## Falta de ajuste y error puro

- Hasta ahora, hemos basado la evaluación de nuestros modelos en el aspecto del gráfico de residuos. Si no hay nada en su distribución que nos haga sospechar una anomalía, consideramos que el modelo es satisfactorio.
- Si nuestro experimento proporciona respuestas duplicadas, podemos utilizarlas para obtener una estimación del error aleatorio. Con esta estimación, tendremos un criterio cuantitativo para juzgar si el modelo elegido es una buena representación de las observaciones, o si necesitamos modificarlo.
- La suma cuadrática residual que deja el modelo puede descomponerse en dos partes: una causada por los errores aleatorios y otra debida a la falta de ajuste del modelo.

15

15

- En cada nivel  $i$  el modelo dejará  $n_i$  residuos, uno por cada respuesta repetida.
- Sumando los cuadrados de todos ellos en todas las repeticiones, obtendremos la suma cuadrática de los residuos en ese nivel. Suponiendo que hay "m" niveles diferentes de la variable X:

**Suma cuadrática de los residuos al nivel  $i$ :**

$$(SQ_r)_i = \sum_j^{n_i} (y_{ij} - \hat{y}_i)^2$$

**Suma cuadrática residual**

$$(SQ_r)_i = \sum_i^m (SQ_r)_i = \sum_i^m \sum_j^{n_i} (y_{ij} - \hat{y}_i)^2$$

$y_{ij}$  = j-ésima respuesta obtenida para el i-ésima ensayo

16

16



- El cuadrado de los residuos puede escribirse como

$$\sum_i^m \sum_j^{n_i} (y_{ij} - \hat{y}_i)^2 = \sum_i^m \sum_j^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_i^m \sum_j^{n_i} (\hat{y}_i - \bar{y}_i)^2$$

- La primera suma del lado derecho no tiene nada que ver con el modelo, y por lo tanto no depende de las estimaciones de  $\hat{y}_i$ , reflejando únicamente la dispersión, en cada nivel  $i$ , de las respuestas repetidas  $y_{ij}$  en torno a sus propias medias  $\bar{y}_i$ .
- Este término, que nos dará una medida del error aleatorio, se denomina **suma cuadrática debida al error puro** ( $SQ_{ep}$ ).
- La segunda suma depende del modelo, y será mayor cuanto más se desvíen las estimaciones para un nivel dado,  $\hat{y}_i$ , si se desvían de la respuesta media correspondiente,  $\bar{y}_i$ . Esto proporciona una medida de la falta de ajuste del modelo a las respuestas observadas, y se denomina **suma al cuadrado por falta de ajuste**,  $SQ_{faj}$

17

17

$$[SQ \text{ residual}] = [SQ \text{ por error puro}] + [SQ \text{ por falta de ajuste}]$$

$$SQ_r = SQ_{ep} + SQ_{faj}$$

Cuando dividamos las sumas cuadráticas por sus respectivos números de grados de libertad, tendremos medias cuadráticas cuyos valores compararemos para evaluar la falta de ajuste del modelo.

$$v_{ep} = (n - m)$$

Donde  $n$  es el número total de observaciones y  $m$  es el número de experimentos únicos.

$$v_{faj} = (m - p)$$

Podemos utilizar la prueba F de la relación  $MQ_{faj}/MQ_{ep}$  para evaluar si el modelo se ajusta bien o no a las observaciones. Valores elevados de esta ratio significarán un ajuste muy pobre, y viceversa.

18

18

Fuente de Variación	Suma Cuadrática	G.L.	Media Cuadrática
Regresión	$\sum_i^m \sum_j^{n_i} (\hat{y}_i - \bar{y})^2$	p-1	$MQ_R = SQ_R / (p-1)$
Residuos	$\sum_i^m \sum_j^{n_i} (y_{ij} - \hat{y}_i)^2$	n-p	$MQ_r = SQ_r / (n-p)$
Falta de Ajuste	$\sum_i^m \sum_j^{n_i} (\hat{y}_i - \bar{y}_i)^2$	m-p	$MQ_{faj} = SQ_{faj} / (m-p)$
Error puro	$\sum_i^m \sum_j^{n_i} (y_{ij} - \bar{y}_i)^2$	n-m	$MQ_{ep} = SQ_{ep} / (n-m)$
Total	$\sum_i^m \sum_j^{n_i} (y_{ij} - \bar{y})^2$	n-1	

$\% \text{ de variación explicada} = \frac{SQ_R}{SQ_T}$      
 $\% \text{ máxima variación explicable} = \frac{SQ_T - SQ_{ep}}{SQ_T}$

19

Estrategia secuencial para lograr condiciones óptimas		
Etapas	Objetivos	Situación
Selección de variables (Factorial fraccionado o de Plackett & Burman)	Identificar las variables más importantes	<ul style="list-style-type: none"> <li>Muchos factores</li> <li>Poco conocimiento del proceso/lejos de las condiciones deseadas u optimizadas</li> </ul>
Otimización (Factorial Completo)	Construir modelos predictivos	<ul style="list-style-type: none"> <li>Pocos factores</li> <li>Dentro de la región óptima</li> </ul>
Validación de las condiciones optimizadas	Confirmar experimentalmente los resultados obtenidos por el análisis de superficie de respuesta	<ul style="list-style-type: none"> <li>Condiciones optimizadas definidas</li> </ul>

20

## Método de superficie de respuesta

Las principales aplicaciones de la MSR son:

- Mapeamiento de una superficie dentro de la región explorada;
- Elección de las condiciones operacionales para obtener una respuesta determinada;
- Búsqueda de las **condiciones óptimas** o, al menos, de las mejores condiciones en la región de interés.

La estrategia de optimización puede resumirse en dos etapas:

- Exploración de experimentos para seleccionar las variables (diseños factoriales).
- Localización/certificación de la subregión óptima mediante diseño experimentales para obtener modelos cuadráticos.

21

21

## Diseño rotacional compuesto central(DCCR)

Son los más adecuados para obtener modelos cuadráticos

$$y = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k \sum_{j>i}^k b_{ij} x_i x_j + \sum_{i=1}^k b_{ii} x_i^2 + \varepsilon$$

Número de factores (k)	2	3	4
$n_{\text{ensayos}}$	4	8	16
Puntos axiales	4	6	8
$\alpha$ (rotación)	1,414	1,682	2
Repeticiones en el punto central	4	4	4
Total de experimentos	12	18	28

22

22

## Puntos Centrales (PC)

- Para poder realizar una inferencia estadística, al menos aproximada, es habitual realizar algunos experimentos en el punto central del espacio experimental.
- De este modo, podremos calcular los residuos y, en consecuencia, el error estándar y, a continuación, las estimaciones por intervalo de los experimentos, etc. También proporciona una medida de la estabilidad del proceso.
- Los ensayos del punto central pueden proporcionar información muy útil sobre el comportamiento de las respuestas entre los niveles asignados inicialmente a los factores, además de mostrar la calidad de la repetibilidad del proceso.
- Comprobar la curvatura.
- La situación ideal es aquella en la que podemos repetir los ensayos en los puntos factoriales y, además, realizar algunas en el punto central. Sin embargo, esto no suele ser posible en ensayos con costos elevados y grandes exigencias de equipamiento y/o tiempo.

23

23

### Ejemplo 02

pH	Temperatura	Actividad Enzimática
-1	-1	200
1	-1	72
-1	1	404
1	1	250
0	0	150
0	0	140
0	0	160

24

24

pH	Temperatura	pH x T	Atividade
-1	-1	1	200
1	-1	-1	72
-1	1	-1	404
1	1	1	250
0	0	0	150
0	0	0	140
0	0	0	160

Ensaio	Calculado	Resíduos	SQR	SQr	Total
1	165,07	34,93	992,25	1220,0	11,76
2	37,07	34,93	25440,25	1220,0	15518,04
3	369,07	34,93	29756,25	1220,0	43026,61
4	215,07	34,93	342,25	1220,0	2854,61
5	196,57	-46,57	0,00	2168,9	2168,90
6	196,57	-56,57	0,00	3200,3	3200,33
7	196,57	-36,57	0,00	1337,5	1337,47
Somatórios			56531	11586,71	68117,71

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$$

	Efeitos	Coefficiente Regressão
Média	196,57	196,57
pH	-141	-70,5
T	191	95,5
pH vs. T	-13	-6,5

$$y = 196,57 - 70,5 x_1 + 95,5 x_2 - 6,5 x_1x_2$$

Média dos PC  
150

Ensaio	Erro Puro	n	p
5	0	7	
6	100	5	
7	100	4	
Soma	200		

25

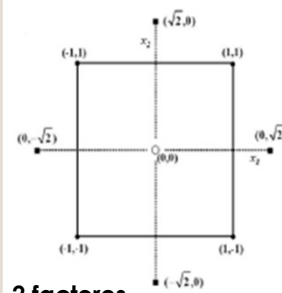
### Tabela ANOVA

Fonte de variação	SQ	G.L	MQ	F Calculado	F Tabelado	Teste
Regressão	56531	3	18843,67	4,88	9,277	Não Significativo
Resíduos	11586,71	3	3862,238			
Falta de Ajuste	11386,71	1	11386,71	113,87	18,513	Significativo
Erro Puro	200	2	100			
Total	68117,71	6				

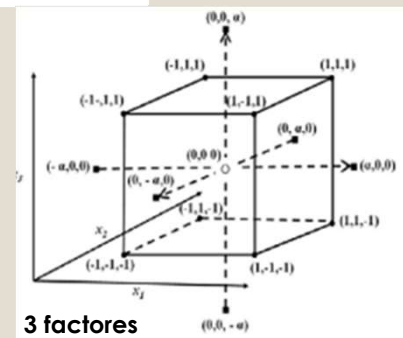
26

Un diseño compuesto central para  $k$  factores consta de tres partes:

1. Una parte denominada factorial (o cúbica), que contiene un total de  $n$  puntos con coordenadas  $-1$  o  $+1$ .
2. Una parte axial (o estrella), formada por  $2k$  puntos con todas las coordenadas cero excepto una, que es igual a un determinado valor  $\alpha$  (o  $-\alpha$ ).
3. Se realizan un total de  $n_c$  ensayos en el punto central, donde todas las coordenadas son 0.



2 factores



3 factores

27

27

- Para llevar a cabo un diseño compuesto central, tenemos que definir cómo será cada una de estas tres partes.
- Los puntos cúbicos son idénticos a los de un diseño factorial completo o, dependiendo del número de factores puede ser fraccional.
- Desde el punto de vista de la resolución, se recomienda utilizar un factorial fraccional de resolución  $V$  (los efectos principales se confunden con interacciones de 4 y los efectos de 2 se confunden con interacciones de 3).
- La elección de resoluciones más pequeñas no es muy trivial y dificulta el análisis de los resultados.
- Las repeticiones en el punto central sirven para dos cosas:
  1. Proporcionar una medida del error puro;
  2. Estabilizar la varianza de la respuesta prevista. Para estabilizar la varianza, es mejor hacer de 2 a 3 si es cúbica.

28

28

## Tipos de Diseños

### Diseño Rotacional

La rotacionalidad permite la homogeneidad de la varianza en todas las direcciones. Para ser rotacional, un diseño cuya porción cúbica sea un factorial completo o un factorial fraccional de resolución V debe tener  $\alpha = \sqrt[4]{n_f}$

### Diseño Box-Behnken ( $3^k$ )

Es un diseño esférico y al menos aproximadamente rotacional, pero sólo se necesitan tres niveles para las variables (-1), (+1) e (0).

### Matriz de Doehlert

Diseño hexagonal con diferente número de niveles para cada factor. Menor número de experimentos en comparación con el DCCR.

29

29

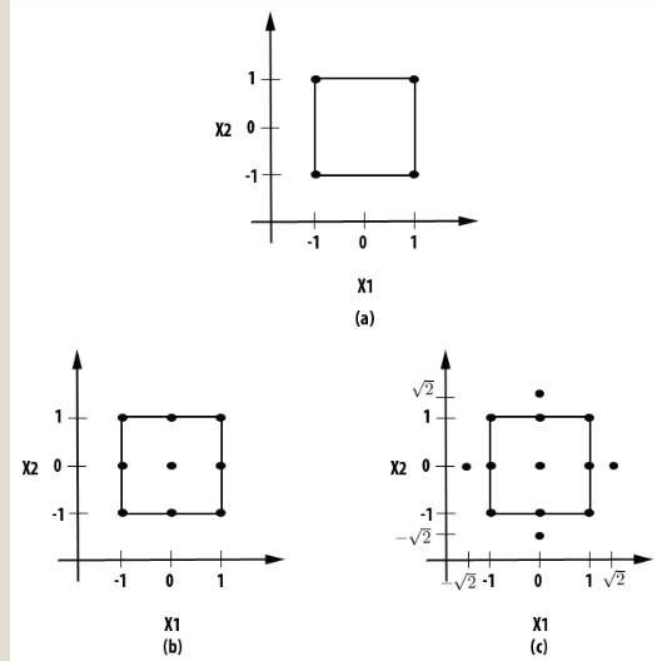
## Diseño Rotacional

- Los experimentos compuestos centrales son los más populares de los diseños experimentales de segundo orden.
- Básicamente, estos experimentos se componen de un punto central, que se ejecutará con réplicas y dará una estimación interna del error puro, y puntos axiales, que determinarán los términos cuadráticos.
- Estos experimentos son de dos niveles totales o factoriales fraccionarios que fueron aumentados con un pequeño número de tratamientos, cuidadosamente escogidos, para permitir la estimativa del modelo de superficie de respuesta de segunda orden.

30

30

- a) Experimento factorial  $2^2$   
 b) Experimento compuesto central con  $\alpha = 1$   
 c) Experimento de composición central con  $\alpha = \sqrt{2}$



31

31

No. Factores	2	3	4	5	6	7	8
Experimento factorial	$2^2$	$2^3$	$2^4$	$2\sqrt{5-1}$	$2\sqrt{6-1}$	$2\sqrt{7-1}$	$2\sqrt{8-2}$
Puntos estrella	4	6	8	10	12	14	16
Puntos de centro	1	1	1	1	1	1	1
$\alpha$	1,414	1,682	2	2	2,378	2,828	3,364
Repeticiones PC	3	3	3	3	3	3	3
Total de ensayos	12	18	28	30	48	82	84

32

32



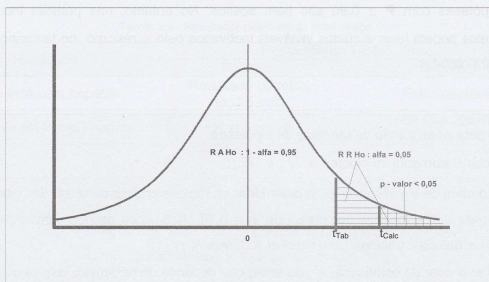
## Probabilidad de significancia, valor p

- Cuando la prueba de hipótesis se realiza en un computador utilizando un programa estadístico, el resultado es el valor p (p-value), el nivel descriptivo o probabilidad de significancia de la prueba, que es la probabilidad de que se produzcan valores de la variable "V" en la prueba que sean más extremos que los obtenidos en la muestra.
- Así, la decisión puede tomarse en función del p-valor: Rechazamos o no  $H_0$ , dependiendo si el p-valor es menor o menor que el nivel de significancia establecido a priori.

33

33

Rejeita-se  $H_0$ , quando  $t_{Calc} \geq t_{Tab}$  ou, equivalentemente, quando p - valor  $\leq \alpha = 0,05$ .



Aquí se muestra un gráfico para una prueba unilateral a la derecha como variable de prueba:

$$V = t \text{ Student}$$

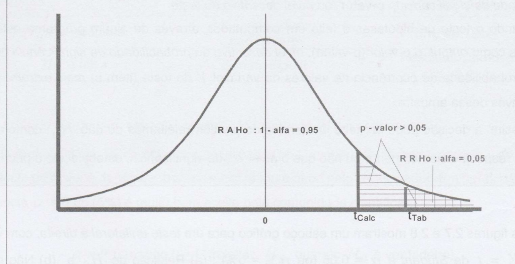
$$\alpha = 0,05$$

- (a) Rechazo de  $H_0$  (Significativo)  
 (b) No rechaza  $H_0$

$\alpha = 0,05$  = nivel de significancia = área de la directa de  $t_{Tab}$

p-valor = área a la directa de  $t_{Calc}$

Não se Rejeita  $H_0$ , quando  $t_{Calc} < t_{Tab}$  ou, equivalentemente, quando p - valor  $> \alpha = 0,05$ .



34

# PRUEBA DE CURVATURA

35

## Ejemplo 3. Diseño factorial 2<sup>2</sup> con tres puntos centrales

Ensayo	t/min	T (°C)	Y (g)
1	70	127,5	54,3
2	80	127,5	60,3
3	70	132,5	64,6
4	80	132,5	68,0
5	75	130	60,3
6	75	130	64,3
7	75	130	62,3

Variables codificadas,  $x_1$  e  $x_2$

$$x_1 = \frac{\text{Tiempo} - PC}{\left(\frac{(+1) 80 - (-1) 70}{2}\right)} = \frac{\text{Tiempo} - 75}{5}$$

$$x_2 = \frac{\text{Temperatura} - PC}{\left(\frac{(+1) 132,5 - (-1) 127,5}{2}\right)} = \frac{\text{Temperatura} - 130}{2,5}$$

**Modelo**

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + \varepsilon$$

**Ajuste de mínimos cuadrados**

$$y = 62,01 + 2,35 x_1 + 4,50 x_2 - 0,65 x_1 x_2$$

36

36

### Inclusión del error a partir de $\sigma$

$$y = 62,01 (\pm 0,76) + 2,35 x_1 (\pm 1,0) + 4,50 (\pm 1,0) x_2 - 0,65 (\pm 1,0) x_1 x_2$$

Los efectos calculados en esta regresión corresponden al doble de los valores de los coeficientes.

$$\text{Efeito (t)} = 4,7 \pm 2,0$$

$$\text{Efeito (T)} = 9,0 \pm 2,0$$

$$\text{Efeito (t x T)} = -1,3 \pm 2,0$$

37

37

### Estimando la curvatura de la superficie, $E_C$

$$E_C = \bar{y}_f - \bar{y}_{PC}$$

$\bar{y}_f$  = Media de los puntos del factorial  $2^2$

$\bar{y}_{PC}$  = Media de los puntos centrales

$$\bar{y}_f = \frac{54,3 + 60,3 + 64,6 + 68}{4} = 61,8$$

$$\bar{y}_{PC} = \frac{60,3 + 64,3 + 62,3}{3} = 62,30$$

$$E_C = 61,8 - 62,30 = -0,50$$

38

38

### Verificación de la Curvatura

Como  $\sigma = 2,0$

$$V(E_C) = V(\bar{y}_f - \bar{y}_{PC}) = V(\bar{y}_f) + V(\bar{y}_{PC}) = \left(\frac{\sigma}{\sqrt{N_f}}\right)^2 + \left(\frac{\sigma}{\sqrt{N_{PC}}}\right)^2 = \left(\frac{2,0}{\sqrt{4}}\right)^2 + \left(\frac{2,0}{\sqrt{3}}\right)^2$$

$$V(E_C) = 2,33 \rightarrow s_C = 1,53$$

Luego,

$$E_C = -0,50 \pm 1,53$$

No hay motivos para cuestionar la adecuación del modelo lineal. Los intervalos de confianza incluyen el cero, la curvatura no es significativa.

39

39

### Ejemplo 4

pH	Temperatura	Actividad Enzimática
-1	-1	200
1	-1	72
-1	1	404
1	1	250
0	0	150
0	0	140
0	0	160
-1,414	0	320
1,414	0	150
0	-1,414	100
0	1,414	300

40

40