

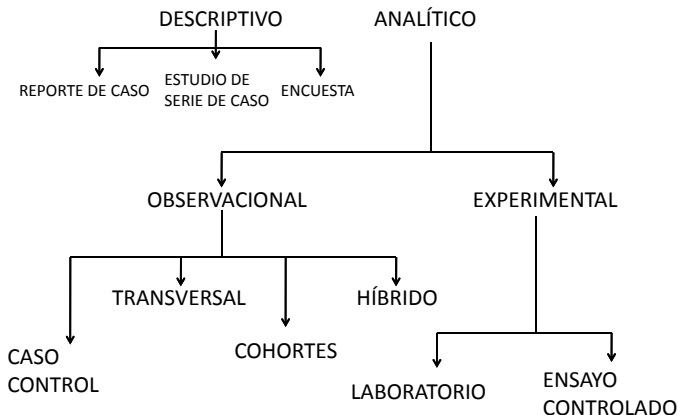
Pruebas de hipótesis

Programa de Maestría en Biotecnología

Daniel Martínez Bello
Universidad de Santander

Noviembre de 2025

Tipos de estudio epidemiológico (Dohoo, Martin, Stryhn, Veterinary Epidemiologic Research, 2004)



- Experimentales: el investigador asigna animales a varios grupos, de acuerdo a factores que el investigador asigna al azar a los animales.
- Observacionales: los animales son asignados a grupos de estudio de acuerdo a características que poseen, pero estos estudios ocurren en condiciones de campo y no es posible asignar animales al azar a grupos de tratamiento, porque el investigador no tiene ningun controls sobre los factores estudiados.

Experimental

- **Laboratorio:** Los individuos son asignados al azar a los tratamientos. Todos los individuos del estudio reciben los mismos cuidados y condiciones de manejo, son uniformes en sus características. Los investigadores controlan completamente todas las condiciones y solo modifican el tratamiento bajo interés.
- **Ensayos controlados:** pruebas clínicas, se tiene una población de sujetos de investigación, que tienen alguna condición especial, sobre los cuales se prueba un tratamiento, los individuos reciben la asignación del tratamiento al azar.

Evaluation of a mycotoxin adsorbent in swine diets containing barley naturally contaminated with *Fusarium* mycotoxins^a

Evaluación de un adsorbente de micotoxinas en dietas para cerdos que contienen cebada contaminada naturalmente con micotoxinas de Fusarium

Avaliação de um adsorvente de micotoxinas em suínos dieta contendo cevada naturalmente contaminado com micotoxinas Fusarium

Changsu Kong, AnSc, PhD; Chan Sol Park, AnSc, MS; Beob Gyun Kim*, AnSc, PhD.

Department of Animal Science and Technology, Konkuk University, Seoul 05029, Republic of Korea.

(Received: September 12, 2014; accepted: March 3, 2016)

doi: 10.17533/udea.rccp.v29n3a02

Summary

Background: contamination of barley with *Fusarium* mycotoxins causes significant economic losses for pork producers, and alleviation with mycotoxin sequestering agents has proven inconsistent. **Objective:** to evaluate a yeast cell wall product in preventing adverse effects of *Fusarium* mycotoxins on growth performance, blood characteristic, and nutrient digestibility in gilts fed diets containing barley naturally contaminated with *Fusarium* mycotoxins. **Methods:** positive and negative controls of corn-soybean meal diets containing 20% control and contaminated barley with *Fusarium* mycotoxins, respectively, were prepared. Two additional diets were prepared by adding 0.2 or 0.4% of a yeast cell wall product to the negative control diet. The experimental diets were fed to pigs with 61.7 Kg initial body weight for 2 weeks. **Results:** pigs fed the negative control diet gained less than those fed the positive control diet ($p<0.05$) from d 0 to 7 and during the overall period, but nutrient digestibility and blood characteristics were not affected by feeding the contaminated diet. Most measurements were not affected by supplementing the yeast cell wall to the negative control diet. **Conclusion:** addition of the yeast cell wall product to negative control diets failed to ameliorate the adverse effects of dietary *Fusarium* mycotoxin on growth performance.

Keywords: *deoxynivalenol, digestion, growth, yeast cell wall.*

Experimental laboratorio

Table 2. Growth performance of pigs fed experimental diets containing barley naturally contaminated with *Fusarium* mycotoxins.

Item ²	Diet ¹				RMSE ³	p-values for contrast ⁴		
	PC	NC	NC1	NC2		PC vs NC	Linear	Quadratic
n	4	4	3	4				
Initial BW, Kg	62.1	62.1	63.4	61.7	-	-	-	-
d 0 to 7								
ADG, g/d	829	261	440	368	194	0.003	0.457	0.381
ADFI, g/d	2,306	1,746	1,751	1,977	364	0.061	0.395	0.676
G:F	0.36	0.13	0.26	0.18	0.10	0.011	0.528	0.168
Final BW, Kg	67.9	63.9	66.5	64.3	1.5	0.006	0.719	0.056
d 7 to 14								
ADG, g/d	832	889	698	825	262	0.765	0.737	0.409
ADFI, g/d	2,388	2,217	1,922	2,216	304	0.449	0.999	0.202
G:F	0.35	0.40	0.36	0.36	0.10	0.576	0.626	0.846
Final BW, Kg	73.7	70.1	71.4	70.1	2.5	0.075	0.978	0.487
d 0 to 14								
ADG, g/d	830	575	569	596	131	0.025	0.823	0.858
ADFI, g/d	2,347	1,981	1,837	2,097	301	0.124	0.602	0.363
G:F	0.36	0.29	0.31	0.28	0.05	0.096	0.605	0.474

¹PC = corn-soybean meal diet containing 20% of control barley; NC = corn-soybean meal diet containing 20% of contaminated barley; NC1 = NC diet supplemented with 0.2% mycotoxin adsorbent; NC2 = NC diet supplemented with 0.4% mycotoxin adsorbent.

²BW = body weight; ADG = average daily gain; ADFI = average daily feed intake; G:F = gain:feed.

³RMSE = root mean square error.

⁴PC vs NC = contrast between the PC diet and the NC diet; Linear = linear effect of the graded concentration of the adsorbent; Quadratic = quadratic effect of the graded concentration of the adsorbent.

Training farmers in rational drug-use improves their management of cattle trypanosomosis: A cluster-randomised trial in south Mali

Delia Grace^{a,b,*}, Thomas Randolph^b,
Oumar Diall^c, Peter-Henning Clausen^a

^a *Institut für Parasitologie und Internationale Tiergesundheit, Freie Universität Berlin,
Königsoweg 67, 14163 Berlin, Germany*

^b *International Livestock Research Institute (ILRI), PO Box 30709, 00100 Nairobi, Kenya*

^c *International Livestock Research Institute (ILRI) Project, International Crops Research
Institute for the Semi-Arid Tropics (ICRISAT-Mali), BP 320 Bamako, Mali*

Received 16 May 2006; received in revised form 11 June 2007; accepted 26 June 2007

Abstract

We carried out a stratified, cluster-randomised, controlled trial in south Mali in 2004 to evaluate the impact of providing information on the diagnosis and treatment of bovine trypanosomosis by farmers. We recruited cattle farmers (444) in 46 villages and used stratified, restricted-randomisation to assign villages to either the test or control group. Farmers in the test group received an information leaflet designed to address gaps in farmer knowledge likely to lead to inadequate treatment; their knowledge was assessed before the intervention, and at 2 weeks and 5 months after the intervention. We assessed the quality of farmer treatments by measuring clinical outcomes in cattle 2 weeks after selection and treatment. As an indicator of herd health, we assessed the mean hematocrit of the village herd before, and 5 months after, the intervention. To account for clustering, we analysed data using generalised estimating equations. Improvements in farmer knowledge of trypanosomosis diagnosis and treatment at 2 weeks and 5 months in the group receiving information were 23% and 14% greater at 2 weeks and 5 months, respectively. In the test group, 84% of farmer treatments were successful, compared to 73% in the control group. Giving rational drug-use information to farmers improved their knowledge and management of trypanosomosis as well as clinical outcomes in cattle they treated and had no discernible negative impacts.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Trypanosomosis; Trypanosomiasis; Rational drug-use; Africa; Cluster-randomised trial; Cattle

Experimental ensayo clínico

Table 2

Changes in primary and secondary outcomes related to bovine trypanosomosis in a comparison of a rational drug-use and a control group in south Mali, April–November 2004

	Test group	<i>n</i>	Control group	<i>n</i>
Primary outcomes: farmer knowledge, cattle successful treatments and herd hematocrit				
Increase in farmer knowledge at 2 weeks (%)	31 ± 21	198	8 ± 20	212
Increase in farmer knowledge at 5 months (%)	19 ± 24	214	5 ± 22	217
Successful treatments of sick cattle (%)	83.7	380	73.1	431
Increase in herd hematocrit at 5 months (%)	2 ± 6	1199	1 ± 6	1294
Secondary outcomes: cattle identified as sick and treated by farmers				
Receive recommended dose trypanocide (%)	59.6	391	29.4	422
Increase in hematocrit after treatment (%)	0.5 ± 4	366	1 ± 4	416
Temperature decrease after treatment (°C)	0.17 ± 0.80	376	0.10 ± 0.76	379
Improvement reported by farmers (%)	95.0	398	89.9	426
Secondary outcomes: cattle trypanosomosis 5 months after intervention				
Herd prevalence of trypanosomosis (%)	5.8	1293	6.9	1356
Reported cases of trypanosomosis per herd	4 ± 8	221	2 ± 4	205
Doses of diminazene given per herd	11 ± 18	221	10 ± 18	205
Doses of isometamidium given per herd	8 ± 19	221	4 ± 14	205

Data are expressed as mean ± standard deviation or percentage.

Observacional

- **Caso control:** individuos con la respuesta de interés son identificados y la historia de exposición de estos casos es contrastada con la historia de exposición de una muestra de individuos (frecuentemente elegida al azar de una fuente definida) de no casos (denominada controles). Se realizan comunmente utilizando bases de datos (retrospectivo), o se puede definir prospectivamente, y en este caso los casos son enrolados en el estudio a medida que van a ocurriendo.
- **Transversal:** una muestra de sujetos de estudio se obtiene y la prevalencia (el número de individuos afectados, sobre el total de individuos) de un evento y la exposición al evento son determinadas.
- **Cohortes:** una muestra de sujetos con diversidad en la exposición al factor (o factores) de interés, o dos o mas grupos definidos por un estado de exposición conocido, son obtenidos, y se determina la presentación de la respuesta en un período de tiempo de seguimiento.

Factors associated with *Mycobacterium avium* subsp. *paratuberculosis* in dairy cows from Northern Antioquia, Colombia[¶]

Factores asociados con el estatus serológico de Mycobacterium avium subsp. paratuberculosis en vacas lecheras del norte de Antioquia, Colombia

Fatores associados ao estado sorológico de Mycobacterium avium subsp. paratuberculosis em soro de gado leiteira no norte de Antioquia, Colombia

Jorge A Fernández-Silva^{1*}, MV, MSP, Dr Med Vet; Nicolás F Ramírez Vásquez², MV, MSc, Dr An Sci; Nathalia M Correa-Valencia², MV, MSc(c), PhD(st).

¹Epidemiología y Medicina Preventiva para la Salud Animal y Humana (EMPreSAH), Escuela de Medicina Veterinaria, Facultad de Ciencias Agrarias, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia.

²Epidemiología y Salud Pública Veterinaria, Centauro Research Group, Escuela de Medicina Veterinaria, Facultad de Ciencias Agrarias, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia.

(Received: March 1, 2016; accepted: September 5, 2016)

doi: 10.17533/udea.rccp.v30n1a06

Abstract

Background: Johne's disease (JD) is a severe enteritis that affects ruminants and has been diagnosed in cattle and sheep in Colombia. However, epidemiological information on the disease in this country is scarce. **Objective:** to identify factors associated with the JD serum enzyme-linked immunosorbent assay (ELISA) status of dairy cows. **Methods:** a cross-sectional study was carried out in 307 asymptomatic adult Holstein dairy cows from 14 herds in nine districts of Belmira and San Pedro de los Milagros municipalities during November, 2007. From 19 to 25 cows, ≥ 2 years of age were randomly selected and blood sampled from every herd. A commercial ELISA kit was used to analyze sera. Information regarding cow related factors (age, farm-born, parity, and daily milk yield) and herd management practices (i.e. herd size, herd average milk production, current presence of symptomatic animals, cattle purchase, own animals grazing in foreign pastures, feeding of calves before weaning, manure spread on pastures, and sighting of birds in feed storing areas) was collected using questionnaires. Descriptive statistics were computed for all variables and a multivariable logistic regression model was constructed ($p < 0.05$). **Results:** ten percent (31/307; 95% CI: 7.0-14.0%) of the animals were positive by ELISA. In 70% (10/14) of the herds, ELISA detected at least one positive animal. Cow and herd factors "parity" and "feeding of calves before weaning" showed

-
- To cite this article: Fernández JA, Ramírez NF, Correa NM. Factors associated with *Mycobacterium avium* subsp. *paratuberculosis* in dairy cows from Northern Antioquia, Colombia. Rev Colomb Cienc Pec 2017; 30: 48-59
 - * Corresponding author: Jorge Arturo Fernández Silva. Escuela de Medicina Veterinaria, Facultad de Ciencias Agrarias, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia. E-mail: jorge.fernandez@udea.edu.co

Table 5. Factors associated to *Mycobacterium avium* subsp. *paratuberculosis* ELISA status (multivariable logistic regression).

Factor	Odds Ratio	Standard error	z	95% CI	P
Parity	1.20	0.12	1.83	0.98-1.47	0.067
Feeding of calves before weaning	0.26	0.13	-2.666	0.09-0.70	0.008
Cons	0.31	0.23	-1.56	0.07-.34	0.120

Wald $\chi^2 = 9.76$; $p = 0.0076$.

Que es una hipótesis

Que es una hipótesis

- Una hipótesis estadística es una afirmación acerca de la distribución de una o varias variables aleatorias.

Que es una hipótesis

- Una hipótesis estadística es una afirmación acerca de la distribución de una o varias variables aleatorias.
- Una hipótesis estadística puede corresponder a una declaración sobre la distribución de probabilidad o a los parámetros de una población que se encuentra bajo estudio.

Que es una hipótesis

- Una hipótesis estadística es una afirmación acerca de la distribución de una o varias variables aleatorias.
- Una hipótesis estadística puede corresponder a una declaración sobre la distribución de probabilidad o a los parámetros de una población que se encuentra bajo estudio.
- La validez de una hipótesis se examina al realizar pruebas de hipótesis sobre observaciones recolectadas en una muestra de la población estudiada.

Que es un experimento?

Que es un experimento?

- "Un experimento es una operación conducida bajo condiciones controladas para descubrir un efecto previamente desconocido, para probar o establecer una hipótesis o para demostrar un ley conocida".

Que es un experimento?

- "Un experimento es una operación conducida bajo condiciones controladas para descubrir un efecto previamente desconocido, para probar o establecer una hipótesis o para demostrar un ley conocida".
- "El objetivo de un experimento es aclarar la relación entre las condiciones controlables y el resultado de un experimento."

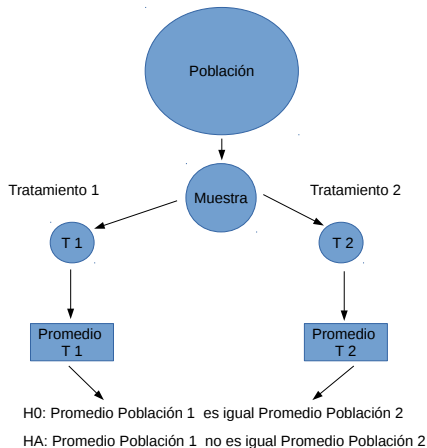
Que es un experimento?

- "Un experimento es una operación conducida bajo condiciones controladas para descubrir un efecto previamente desconocido, para probar o establecer una hipótesis o para demostrar un ley conocida".
- "El objetivo de un experimento es aclarar la relación entre las condiciones controlables y el resultado de un experimento."
- "Un análisis experimental se realiza sobre observaciones que son afectadas no solamente por las condiciones controladas, sino tambien por condiciones no controladas y errores de medida."

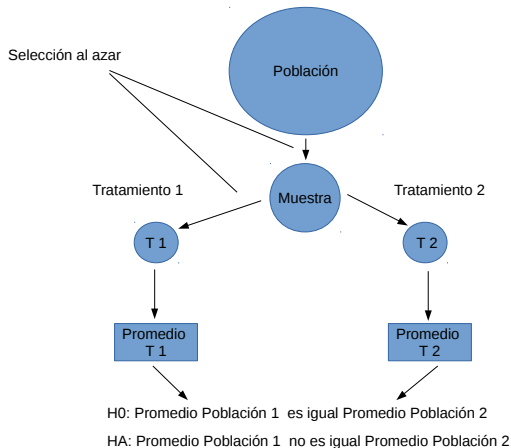
Que es un experimento?

- "Un experimento es una operación conducida bajo condiciones controladas para descubrir un efecto previamente desconocido, para probar o establecer una hipótesis o para demostrar un ley conocida".
- "El objetivo de un experimento es aclarar la relación entre las condiciones controlables y el resultado de un experimento."
- "Un análisis experimental se realiza sobre observaciones que son afectadas no solamente por las condiciones controladas, sino tambien por condiciones no controladas y errores de medida."
- (The Concise Encyclopedia of Statistics, 2008).

Inferencia estadística



Inferencia estadística



No money, no honey!!
No muestra al azar, no inferencia

Pregunta científica:

- Un población produce mas que otra población?

Pregunta científica:

- Un población produce mas que otra población?
- Una población gasta mas que otra población?

Pregunta científica:

- Un población produce mas que otra población?
- Una población gasta mas que otra población?
- Una población tiene mas riesgo que otra población?

Pregunta científica:

- Un población produce mas que otra población?
- Una población gasta mas que otra población?
- Una población tiene mas riesgo que otra población?
- Las diferencias que hay entre diferentes poblaciones sometidas a diferentes tratamientos?

Pruebas de Hipótesis

Parámetro				Estadístico
μ	←	Media	→	\bar{x}
σ	←	Desviación estándar	→	s
π	←	Proporción	→	p
$\frac{\pi}{1-\pi}$	←	Odds	→	$\frac{p}{1-p}$
ρ	←	Correlación	→	r
N	←	Tamaño	→	n

α alfa

β beta

γ gama

δ delta

τ tau

ϵ epsilon

σ sigma

ϕ fi

χ chi

λ lambda

μ mu

ω omega

θ teta

ρ ro

ν nu

η eta

κ kappa

υ upsilon

ξ xi

ψ psi

ι iota

ζ zeta

- Epicardial Adipose Tissue and Insulin Resistance in Obese Subjects. Gianluca Iacobellis and Frida Leonetti, (J Clin Endocrinol Metab 90: 6300 – 6302, 2005)
- *"in this study we evaluated whether epicardial adipose tissue is related to insulin sensitivity and glucose metabolism in obese subjects."*

- A short-term circuit resistance programme reduced epicardial fat in obese aged women. Miguel Ángel Rosety, María Teresa Pery, María Antonia Rodríguez-Pareja, Antonio Díaz, Jesús Rosety, Natalia García, Francisco Brenes-Martín, Manuel Rosety Rodríguez, Rocío Toro, Francisco Javier Ordóñez and Ignacio Rosety, Nutr Hosp. 2015;32(5):2193-2197
- *"this study was conducted to ascertain the effects of resistance circuit training on EAT in obese aged women. A secondary objective was to assess muscle damage induced by supervised resistance training to confirm the intervention program was effective and safe."*

Pasos para formular una prueba de hipótesis

- 1 Definir el problema o la pregunta científica.
- 2 Definir la hipótesis nula H_0 y la hipótesis alterna H_A .
- 3 Definir el nivel de significancia
- 4 Definir la regla de decisión.
- 5 Calcular un estadístico apropiado basado en los datos y calcular el valor-p.
- 6 Comparar el valor-p con el nivel de significancia, y aceptar o rechazar la hipótesis nula.
- 7 Concluir.

Pruebas de Hipótesis

Tipos de problemas en los cuales se pueden aplicar pruebas de hipótesis

Segun la cantidad de poblaciones

- 1 Estimación del promedio de una población.
- 2 Comparación del promedio de dos poblaciones.
- 3 Comparación de varianzas de dos poblaciones.
- 4 Comparación del promedio de tres o mas poblaciones.
- 5 Comparación de proporciones de dos poblaciones.
- 6 Comparación de riesgos de dos poblaciones
- 7 Comparación de odds de dos poblaciones

Según si se asume una distribución de probabilidad para los resultados de los datos

- 1 Pruebas paramétricas.
- 2 Pruebas no paramétricas.

Tipos de pruebas de hipótesis

Estimación del promedio de una población.

❶ Hipótesis a dos colas.

$$H_0 : \mu = \mu_0$$

$$H_A : \mu \neq \mu_0$$

❷ Hipótesis a una cola.

$$H_0 : \mu > \mu_0$$

$$H_A : \mu \leq \mu_0$$

$$H_0 : \mu < \mu_0$$

$$H_A : \mu \geq \mu_0$$

Comparación de promedios de dos poblaciones independientes

1 Hipótesis a dos colas.

$$H_0 : \mu_{Tratamiento} = \mu_{Control}$$

$$H_A : \mu_{Tratamiento} \neq \mu_{Control}$$

Pruebas de Hipótesis

Errores que ocurren cuando se prueba una hipótesis

Definiendo el nivel de significancia

Reportado por el test Estadístico	Resultado Final	
	Situación cierta pero desconocida H_0 es cierta	H_0 es falsa
Hipótesis nula H_0 se rechaza	Error Tipo I α	
Hipótesis nula H_0 se acepta		Error Tipo II β

Comparación del promedio de dos poblaciones

Comparación de la media de dos poblaciones independientes

Reglas de Decisión

$H_0 : \mu_1 = \mu_2$; si $|t^*| \leq t(1 - \alpha/2, n_1 + n_1 - 2)$ Concluir H_0

$H_A : \mu_1 \neq \mu_2$; si $|t^*| > t(1 - \alpha/2, n_1 + n_1 - 2)$ Concluir H_A

Comparación del promedio de dos poblaciones

Comparación de la media de dos poblaciones independientes

Reglas de Decisión

Decido $H_0 : \mu_1 = \mu_2$; si valor-p es mayor que α

Decido $H_A : \mu_1 \neq \mu_2$; si valor-p es menor que α

Comparación del promedio de dos poblaciones

Estadísticos para comparación de las medias de dos poblaciones independientes

Cuando las varianzas de ambas poblaciones son iguales el valor que aparece como t^* se extrae del estadístico,

$$t^* = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Dado que s_1^2 y s_2^2 se asumen iguales, se calcula una varianza de la muestra conjunta denotada como s_p^2 ,

$$t^* = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Comparación del promedio de dos poblaciones

Estadísticos para comparación de las medias de dos poblaciones independientes

Cuando las varianzas de ambas poblaciones son diferentes el valor que aparece como t^* se extrae del estadístico,

$$t^* = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Dado que s_1^2 y s_2^2 son diferentes, se requiere calcular los grados de libertad para buscar en la tabla de t.

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Comparación del promedio de dos poblaciones

Prueba de hipótesis, comparación del promedio de dos poblaciones independientes

Cuando las varianzas σ_1^2 y σ_2^2 desconocidas y diferentes.

$$\begin{aligned} H_0 : \mu_1 &= \mu_2 \\ H_a : \mu_1 &\neq \mu_2 \end{aligned} \quad \begin{aligned} t^* &= \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_\nu \\ \nu &= \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2 - 1}} \end{aligned} \quad \begin{aligned} |t^*| &\leq t_{1-\alpha/2, \nu}, & H_0 \\ |t^*| &> t_{1-\alpha/2, \nu}, & H_a \end{aligned}$$

Comparación del promedio de dos poblaciones

Prueba de hipótesis, comparación del promedio de dos poblaciones independientes

Cuando las varianzas σ_1^2 y σ_2^2 desconocidas y iguales.

$$\begin{array}{ll} H_0 : \mu_1 = \mu_2 & t^* = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(\nu)} \\ H_a : \mu_1 \neq \mu_2 & s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad \begin{array}{l} |t^*| \leq t_{1-\alpha/2, \nu}, H_0 \\ |t^*| > t_{1-\alpha/2, \nu}, H_a \end{array} \\ & \nu = n_1 + n_2 - 2 \end{array}$$

Comparación del promedio de dos poblaciones

Intervalos de confianza, comparación del promedio de dos poblaciones independientes

Cuando las varianzas σ_1^2 y σ_2^2 conocidas.

Parámetro	Estadístico	Intervalo de confianza
$\mu_1 - \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$\bar{x}_1 - \bar{x}_2 \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Comparación del promedio de dos poblaciones

Intervalo de confianza, comparación del promedio de dos poblaciones independientes

Cuando las varianzas σ_1^2 y σ_2^2 desconocidas y diferentes.

Parámetro	Estadístico	Intervalo de confianza
$\mu_1 - \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{(\alpha/2, f)}$	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, f} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
	$f = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2 - 1}}$	

Comparación del promedio de dos poblaciones

Intervalos de confianza, comparación del promedio de dos poblaciones independientes

Cuando las varianzas σ_1^2 y σ_2^2 desconocidas y iguales.

Parámetro	Estadístico	Intervalo de confianza
$\mu_1 - \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(\alpha/2, g)}$	$\bar{x}_1 - \bar{x}_2 \pm t_{(\alpha/2, g)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
	$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	
	$g = n_1 + n_2 - 2$	

Tabla de contingencia dos x dos

Tablas de contingencia como forma de expresar la relación entre dos o mas variables discretas o categóricas.

- Tablas de contingencia: es una tabla que presenta en cada celda la frecuencia de una respuesta.
- Tabla de dos x dos: Es una tabla de contingencia que clasifica cruzadamente dos variables.

Test de independencia de Chi cuadrado en tablas de contingencia 2×2 .

- Donde, una forma resumida de presentar la anterior tabla es

	Variable dos		
	Si	No	
Variable uno			
Si	n_{11}	n_{12}	n_{1+}
No	n_{21}	n_{22}	n_{2+}
	n_{+1}	n_{+2}	n_{++}

Test de independencia de Chi cuadrado en tablas de contingencia 2×2 .

- Hipótesis estadística de asociación

$$H_0 : \pi_{ij} = \pi_{i+}\pi_{+j}$$

$$H_A : \pi_{ij} \neq \pi_{i+}\pi_{+j}$$

- Se define el nivel de significancia, $\alpha = 0.05$.
- Regla de decisión (grados de libertad = $(I - 1)(J - 1)$), forma difícil

$$H_0 : \pi_{ij} = \pi_{i+}\pi_{+j}, \text{ si } \chi^2_* \leq \chi^2(1 - \alpha/2, \text{grados de libertad})$$

$$H_A : \pi_{ij} \neq \pi_{i+}\pi_{+j}, \text{ si } \chi^2_* > \chi^2(1 - \alpha/2, \text{grados de libertad})$$

- Regla de decisión fácil: si valor-p es mayor que α decido la hipótesis nula, si valor-p es menor que α decido la hipótesis alterna.

Test de independencia de Chi Cuadrado en tablas de contingencia 2×2 .

- Y el estadístico es

$$\chi^2_* = \sum_{i=1}^I \sum_{j=1}^J \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

- donde O_{ij} son los conteos observados y E_{ij} son los conteos esperados.

Tabla de contingencia dos x dos

Comparando proporciones en tablas de contingencia.

- La **Diferencia de proporciones** (**Diferencia de riesgo**)
- Se calculan primero las proporciones para cada grupo

$$P_1 = \frac{X_1}{n_1}; P_2 = \frac{X_2}{n_2}$$

- El estimador de la diferencia de proporciones es

$$\text{Diferencia de proporciones} = P_1 - P_2$$

Tabla de contingencia dos x dos

Prueba de hipótesis e intervalo de confianza de 95% para la diferencia de proporciones.

Para probar hipótesis o construir intervalos de confianza, se puede usar el siguiente estadístico (para muestras grandes):

$$Z^* = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \sim \text{Normal}(0, 1)$$

El intervalo de confianza del $(1-\alpha)\%$

$$(p_1 - p_2) \pm Z_{\alpha/2} * \sqrt{\text{Var}(p_1 - p_2)}$$
$$\text{Var}(p_1 - p_2) = \frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}$$

Tabla de contingencia dos x dos

Prueba de hipótesis e intervalo de confianza de 95% para el riesgo relativo.

Hipótesis estadística para el logaritmo del riesgo relativo, a dos colas:

$$H_0 : \log \left(\frac{\pi_1}{\pi_2} \right) = 0$$

$$H_A : \log \left(\frac{\pi_1}{\pi_2} \right) \neq 0$$

Tabla de contingencia dos x dos

Prueba de hipótesis e intervalo de confianza de 95% para el riesgo relativo.

Para probar hipótesis o construir intervalos de confianza, se puede usar el siguiente estadístico (para muestras grandes):

$$Z^* = \frac{\log\left(\frac{p_1}{p_2}\right)}{\sqrt{\frac{(1-p_1)}{n_1 p_1} + \frac{(1-p_2)}{n_2 p_2}}} \sim \text{Normal}(0, 1)$$

El intervalo de confianza del $(1-\alpha)\%$

$$\log\left(\frac{p_1}{p_2}\right) \pm Z_{\alpha/2} * \sqrt{\text{Var}\left[\log\left(\frac{p_1}{p_2}\right)\right]}$$
$$\text{Var}\left[\log\left(\frac{p_1}{p_2}\right)\right] = \frac{(1-p_1)}{n_1 p_1} + \frac{(1-p_2)}{n_2 p_2}$$

Tabla de contingencia dos x dos

Comparando proporciones en tablas de contingencia.

- La razón de Odds

$$P_1 = \frac{X_1}{n_1}$$

$$P_2 = \frac{X_2}{n_2}$$

- El estimador de la razón de Odds es

$$\text{Razon de Odds} = \frac{\frac{P_1}{1-P_1}}{\frac{P_2}{1-P_2}}$$

Tabla de contingencia dos x dos

Prueba de hipótesis e intervalo de confianza de 95% para la razón de odds.

Hipótesis estadística para el logaritmo de la razón de odds, a dos colas:

$$H_0 : \log \left[\frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)} \right] = 0$$

$$H_A : \log \left[\frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)} \right] \neq 0$$

Tabla de contingencia dos x dos

Prueba de hipótesis e intervalo de confianza de 95% para la razón de odds.

Para probar hipótesis o construir intervalos de confianza, se puede usar el siguiente estadístico (para muestras grandes):

$$Z^* = \frac{\log \left[\frac{p_1/(1-p_1)}{p_2/(1-p_2)} \right]}{\sqrt{\frac{1}{n_1(1-p_1)} + \frac{1}{n_1 p_1} + \frac{1}{n_2 p_2} + \frac{1}{n_2(1-p_2)}}} \sim \text{Normal}(0, 1)$$

El intervalo de confianza del $(1-\alpha)\%$

$$\log \left[\frac{p_1/(1-p_1)}{p_2/(1-p_2)} \right] \pm Z_{\alpha/2} * \sqrt{\text{Var} \left(\log \left[\frac{p_1/(1-p_1)}{p_2/(1-p_2)} \right] \right)}$$
$$\text{Var} \left(\log \left[\frac{p_1/(1-p_1)}{p_2/(1-p_2)} \right] \right) = \frac{1}{n_1(1-p_1)} + \frac{1}{n_1 p_1} + \frac{1}{n_2 p_2} + \frac{1}{n_2(1-p_2)}$$

Correlación lineal

Una medida de asociación entre dos variables continuas relacionadas se conoce como coeficiente de correlación lineal

- El coeficiente de correlación para una muestra se denota con la letra latina r .
- El calculo del coeficiente de correlación se hace con la siguiente formula

$$r^2 = \frac{(SC_{XY})^2}{(SC_{XX})(SC_{YY})}$$

- El coeficiente de correlación se obtiene como la raíz cuadrada de r^2 ($r = \sqrt{r^2}$), y r^2 se conoce como el coeficiente de correlación.

Correlación lineal

Una medida de asociación entre dos variables continuas relacionadas se conoce como coeficiente de correlación lineal

- SC significa suma de cuadrados, SC_{XX} denota la suma de los cuadrados de X, SC_{YY} denota la suma de los cuadrados de Y, y SC_{XY} denota la suma de X y Y,

$$SC_{XX} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$SC_{YY} = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

$$SC_{XY} = \sum x_i y_i - \frac{\sum y_i \sum x_i}{n}$$

Hipótesis de la correlación lineal entre dos poblaciones

Con las sumas de cuadrados se estima el coeficiente de correlación de la muestra denotado como r :

$$\frac{SC_{xy}}{\sqrt{SC_{xx}SC_{yy}}} = r$$

Utilizando el coeficiente de correlación lineal de la muestra se calcula el estadístico t^* . Para calcular el estadístico se requiere el error estándar s_r del coeficiente de correlación :

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

Hipótesis de la correlación lineal entre dos poblaciones

La hipótesis de correlación lineal

$$H_0 : \rho = 0$$

$$H_A : \rho \neq 0$$

Regla de decisión (forma larga)

- Área crítica: si $|t^*| \leq t_{(1-\alpha/2, \nu)}$, entonces H_0 .
- Área crítica: si $|t^*| > t_{(1-\alpha/2, \nu)}$, entonces H_A .

Regla de decisión (forma corta): si valor-p es mayor que α decido H_0 , si valor-p es menor que α decido H_A .

Se calcula el estadístico t^* con la siguiente fórmula:

$$t^* = \frac{r - 0}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$