

**UNIVERSIDAD PRIVADA DE TACNA**  
**FACULTAD DE INGENIERÍA**  
**ESCUELA DE INGENIERÍA DE SISTEMAS**



## **Pruebas no parametricas - Unidad 2**

**Que se presenta para el curso:**  
**“ESTADÍSTICA INFERENCIAL Y ANÁLISIS DE DATOS”**

**Docente:**  
**ING. LUIS A. FERNÁNDEZ VIZCARRA**

**Autor(es):**

<b>Vargas Gutierrez, Angel Jose</b>	<b>(2020066922)</b>
<b>Mamani Valdvia, Elvis</b>	<b>(2020068763)</b>

**TACNA – PERÚ**  
**2025**

# INTRODUCCIÓN

## INTRODUCCIÓN

La institución sospecha que el nivel de rendimiento difiere significativamente entre dos grupos identificados en el conjunto de datos. El comité académico desea saber si existe evidencia estadística de asociación entre variables categóricas y la clasificación de rendimiento (Alto y Bajo). Con base en una muestra de 480 observaciones, se procede a contrastar esta hipótesis con un nivel de significación  $\alpha = 0.05$ .

## 1. PRUEBA CHI-CUADRADO

### Hipótesis

- **Hipótesis Nula ( $H_0$ ):** Las variables categóricas son independientes del grupo de rendimiento al que pertenecen. Es decir, no existe una asociación significativa entre estas dos variables.
- **Hipótesis Alternativa ( $H_1$ ):** Las variables categóricas están asociadas al grupo de rendimiento al que pertenecen. Es decir, existe una relación significativa entre estas variables.

Ejecución en SPSS

Resumen de procesamiento de casos

		Casos					
		Válido		Perdidos		Total	
		N	Porcentaje	N	Porcentaje	N	Porcentaje
puntaje	alto	251	100,0%	0	0,0%	251	100,0%
	bajo	229	100,0%	0	0,0%	229	100,0%

id\*grupo tabulación cruzada

Recuento

grupo                  Total

alto    bajo

id        1        1        0        1

2	1	0	1
3	1	0	1
4	1	0	1
5	1	0	1
6	1	0	1
7	1	0	1
8	1	0	1
9	1	0	1
10	1	0	1
11	1	0	1
12	1	0	1
13	1	0	1
14	1	0	1
15	1	0	1
16	1	0	1
17	1	0	1
18	1	0	1
19	1	0	1
20	1	0	1
21	1	0	1
22	1	0	1

23	1	0	1
24	1	0	1
25	1	0	1
26	1	0	1
27	1	0	1
28	1	0	1
29	1	0	1
30	1	0	1
31	1	0	1
32	1	0	1
33	1	0	1
34	1	0	1
35	1	0	1
36	1	0	1
37	1	0	1
38	1	0	1
39	1	0	1
40	1	0	1
41	1	0	1
42	1	0	1
43	1	0	1

44	1	0	1
45	1	0	1
46	1	0	1
47	1	0	1
48	1	0	1
49	1	0	1
50	1	0	1
51	1	0	1
52	1	0	1
53	1	0	1
54	1	0	1
55	1	0	1
56	1	0	1
57	1	0	1
58	1	0	1
59	1	0	1
60	1	0	1
61	1	0	1
62	1	0	1
63	1	0	1
64	1	0	1

65	1	0	1
66	1	0	1
67	1	0	1
68	1	0	1
69	1	0	1
70	1	0	1
71	1	0	1
72	1	0	1
73	1	0	1
74	1	0	1
75	1	0	1
76	1	0	1
77	1	0	1
78	1	0	1
79	1	0	1
80	1	0	1
81	1	0	1
82	1	0	1
83	1	0	1
84	1	0	1
85	1	0	1

86	1	0	1
87	1	0	1
88	1	0	1
89	1	0	1
90	1	0	1
91	1	0	1
92	1	0	1
93	1	0	1
94	1	0	1
95	1	0	1
96	1	0	1
97	1	0	1
98	1	0	1
99	1	0	1
100	1	0	1
101	1	0	1
102	1	0	1
103	1	0	1
104	1	0	1
105	1	0	1
106	1	0	1

107	1	0	1
108	1	0	1
109	1	0	1
110	1	0	1
111	1	0	1
112	1	0	1
113	1	0	1
114	1	0	1
115	1	0	1
116	1	0	1
117	1	0	1
118	1	0	1
119	1	0	1
120	1	0	1
121	1	0	1
122	1	0	1
123	1	0	1
124	1	0	1
125	1	0	1
126	1	0	1
127	1	0	1

128	1	0	1
129	1	0	1
130	1	0	1
131	1	0	1
132	1	0	1
133	1	0	1
134	1	0	1
135	1	0	1
136	1	0	1
137	1	0	1
138	1	0	1
139	1	0	1
140	1	0	1
141	1	0	1
142	1	0	1
143	1	0	1
144	1	0	1
145	1	0	1
146	1	0	1
147	1	0	1
148	1	0	1

149	1	0	1
150	1	0	1
151	1	0	1
152	1	0	1
153	1	0	1
154	1	0	1
155	1	0	1
156	1	0	1
157	1	0	1
158	1	0	1
159	1	0	1
160	1	0	1
161	1	0	1
162	1	0	1
163	1	0	1
164	1	0	1
165	1	0	1
166	1	0	1
167	1	0	1
168	1	0	1
169	1	0	1

170	1	0	1
171	1	0	1
172	1	0	1
173	1	0	1
174	1	0	1
175	1	0	1
176	1	0	1
177	1	0	1
178	1	0	1
179	1	0	1
180	1	0	1
181	1	0	1
182	1	0	1
183	1	0	1
184	1	0	1
185	1	0	1
186	1	0	1
187	1	0	1
188	1	0	1
189	1	0	1
190	1	0	1

191	1	0	1
192	1	0	1
193	1	0	1
194	1	0	1
195	1	0	1
196	1	0	1
197	1	0	1
198	1	0	1
199	1	0	1
200	1	0	1
201	1	0	1
202	1	0	1
203	1	0	1
204	1	0	1
205	1	0	1
206	1	0	1
207	1	0	1
208	1	0	1
209	1	0	1
210	1	0	1
211	1	0	1

212	1	0	1
213	1	0	1
214	1	0	1
215	1	0	1
216	1	0	1
217	1	0	1
218	1	0	1
219	1	0	1
220	1	0	1
221	1	0	1
222	1	0	1
223	1	0	1
224	1	0	1
225	1	0	1
226	1	0	1
227	1	0	1
228	1	0	1
229	1	0	1
230	1	0	1
231	1	0	1
232	1	0	1

233	1	0	1
234	1	0	1
235	1	0	1
236	1	0	1
237	1	0	1
238	1	0	1
239	1	0	1
240	1	0	1
241	0	1	1
242	0	1	1
243	0	1	1
244	0	1	1
245	0	1	1
246	0	1	1
247	0	1	1
248	0	1	1
249	0	1	1
250	0	1	1
251	0	1	1
252	0	1	1
253	0	1	1

254	0	1	1
255	0	1	1
256	0	1	1
257	0	1	1
258	0	1	1
259	0	1	1
260	0	1	1
261	0	1	1
262	0	1	1
263	0	1	1
264	0	1	1
265	0	1	1
266	0	1	1
267	0	1	1
268	0	1	1
269	0	1	1
270	0	1	1
271	0	1	1
272	0	1	1
273	0	1	1
274	0	1	1

275	0	1	1
276	0	1	1
277	0	1	1
278	0	1	1
279	0	1	1
280	0	1	1
281	0	1	1
282	0	1	1
283	0	1	1
284	0	1	1
285	0	1	1
286	0	1	1
287	0	1	1
288	0	1	1
289	0	1	1
290	0	1	1
291	0	1	1
292	0	1	1
293	0	1	1
294	0	1	1
295	0	1	1

296	0	1	1
297	0	1	1
298	0	1	1
299	0	1	1
300	0	1	1
301	0	1	1
302	0	1	1
303	0	1	1
304	0	1	1
305	0	1	1
306	0	1	1
307	0	1	1
308	0	1	1
309	0	1	1
310	0	1	1
311	0	1	1
312	0	1	1
313	0	1	1
314	0	1	1
315	0	1	1
316	0	1	1

317	0	1	1
318	0	1	1
319	0	1	1
320	0	1	1
321	0	1	1
322	0	1	1
323	0	1	1
324	0	1	1
325	0	1	1
326	0	1	1
327	0	1	1
328	0	1	1
329	0	1	1
330	0	1	1
331	0	1	1
332	0	1	1
333	0	1	1
334	0	1	1
335	0	1	1
336	0	1	1
337	0	1	1

338	0	1	1
339	0	1	1
340	0	1	1
341	0	1	1
342	0	1	1
343	0	1	1
344	0	1	1
345	0	1	1
346	0	1	1
347	0	1	1
348	0	1	1
349	0	1	1
350	0	1	1
351	0	1	1
352	0	1	1
353	0	1	1
354	0	1	1
355	0	1	1
356	0	1	1
357	0	1	1
358	0	1	1

359	0	1	1
360	0	1	1
361	0	1	1
362	0	1	1
363	0	1	1
364	0	1	1
365	0	1	1
366	0	1	1
367	0	1	1
368	0	1	1
369	0	1	1
370	0	1	1
371	0	1	1
372	0	1	1
373	0	1	1
374	0	1	1
375	0	1	1
376	0	1	1
377	0	1	1
378	0	1	1
379	0	1	1

380	0	1	1
381	0	1	1
382	0	1	1
383	0	1	1
384	0	1	1
385	0	1	1
386	0	1	1
387	0	1	1
388	0	1	1
389	0	1	1
390	0	1	1
391	0	1	1
392	0	1	1
393	0	1	1
394	0	1	1
395	0	1	1
396	0	1	1
397	0	1	1
398	0	1	1
399	0	1	1
400	0	1	1

401	0	1	1
402	0	1	1
403	0	1	1
404	0	1	1
405	0	1	1
406	0	1	1
407	0	1	1
408	0	1	1
409	0	1	1
410	0	1	1
411	0	1	1
412	0	1	1
413	0	1	1
414	0	1	1
415	0	1	1
416	0	1	1
417	0	1	1
418	0	1	1
419	0	1	1
420	0	1	1
421	0	1	1

422	0	1	1
423	0	1	1
424	0	1	1
425	0	1	1
426	0	1	1
427	0	1	1
428	0	1	1
429	0	1	1
430	0	1	1
431	0	1	1
432	0	1	1
433	0	1	1
434	0	1	1
435	0	1	1
436	0	1	1
437	0	1	1
438	0	1	1
439	0	1	1
440	0	1	1
441	0	1	1
442	0	1	1

443	0	1	1
444	0	1	1
445	0	1	1
446	0	1	1
447	0	1	1
448	0	1	1
449	0	1	1
450	0	1	1
451	0	1	1
452	0	1	1
453	1	0	1
454	1	0	1
455	1	0	1
456	1	0	1
457	1	0	1
458	1	0	1
459	1	0	1
460	1	0	1
461	1	0	1
462	1	0	1
463	1	0	1

464	0	1	1
465	0	1	1
466	0	1	1
467	0	1	1
468	0	1	1
469	0	1	1
470	0	1	1
471	0	1	1
472	0	1	1
473	0	1	1
474	0	1	1
475	0	1	1
476	0	1	1
477	0	1	1
478	0	1	1
479	0	1	1
480	0	1	1
Total	251	229	480

### Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)
Chi-cuadrado de Pearson	480,000 <sup>a</sup>	479	,479
Razón de verosimilitud	664,413	479	,000
Asociación lineal por lineal	302,678	1	,000
N de casos válidos	480		

a. 960 casillas (100,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es ,48.

#### Estadísticos de la Prueba:

- **Estadístico Chi-cuadrado ( $\chi^2$ ):** 480.000
  - Este valor se obtiene del análisis realizado en SPSS
- **Grados de libertad (gl):** 479
- **Valor p (Significación):** 0.479
  - El valor p es el resultado del test de Chi-cuadrado, que nos indica la significación de la asociación entre las dos variables

## DISCUSIÓN (Toma de Decisión)

#### Criterio de Decisión:

- Si  $p \leq 0.05$ : Se rechaza  $H_0$
- Si  $p > 0.05$ : No se rechaza  $H_0$

Dado que el valor p obtenido (0.479) es **considerablemente mayor** que el nivel de significancia  $\alpha = 0.05$ , **NO se rechaza la hipótesis nula ( $H_0$ )**. Esto indica que **NO hay una asociación estadísticamente significativa** entre la variable categórica analizada y el grupo de rendimiento al que pertenecen. En otras palabras, la variable es independiente del grupo de rendimiento.

## CONCLUSIÓN (Interpretación)

Se concluye que **NO existe una asociación estadísticamente significativa** entre la variable categórica y la pertenencia al grupo de Alto o Bajo Rendimiento ( $\chi^2 = 480.000$ ,  $gl = 479$ ,  $p =$

0.479). La distribución de la variable categórica es aleatoria e independiente entre los grupos de rendimiento, lo que sugiere que ambas variables no están relacionadas.

**Nota Importante:** El análisis muestra que 960 casillas (100%) tienen frecuencias esperadas menores a 5, lo que indica que la prueba Chi-cuadrado **no es apropiada** para estos datos debido a que no se cumplen los supuestos necesarios para su aplicación válida.

## Hipótesis

- **Hipótesis Nula ( $H_0$ ):** La proporción de casos en el grupo de Alto Rendimiento es igual al 50% ( $p = 0.50$ ).
- **Hipótesis Alternativa ( $H_1$ ):** La proporción de casos en el grupo de Alto Rendimiento es diferente al 50% ( $p \neq 0.50$ ).

### → Pruebas NPar

Estadísticos descriptivos

	N	Media	Desviación estándar	Mínimo	Máximo	Percentiles		
						Percentil 25	Percentil 50 (Mediana)	75°
grupo	480	1,48	,500	1	2	1,00	1,00	2,00

Prueba binomial

	Categoría	N	Prop. observada	Prop. de prueba	Significación exacta (bilateral)
grupo	Grupo 1 alto	251	,52	,50	,338
	Grupo 2 bajo	229	,48		
	Total	480	1,00		

## DISCUSIÓN

### Estadísticos de la Prueba:

- **Proporción observada (Grupo Alto):** 0.52 (52%)

- **Proporción observada (Grupo Bajo):** 0.48 (48%)
- **Proporción de prueba:** 0.50 (50%)
- **Significación exacta (bilateral):**  $p = 0.338$
- **Tamaño de muestra:**  $N = 480$

#### **Criterio de Decisión:**

- Si  $p \leq 0.05$ : Se rechaza  $H_0$
- Si  $p > 0.05$ : No se rechaza  $H_0$

Dado que el valor  $p$  obtenido (0.338) es **mayor** que el nivel de significancia  $\alpha = 0.05$ , **NO se rechaza la hipótesis nula ( $H_0$ )**. Esto indica que la proporción observada de casos en el grupo de Alto Rendimiento (52%) **no difiere significativamente** de la proporción esperada (50%).

---

## **CONCLUSIÓN (Interpretación)**

Se concluye que la proporción de casos en el grupo de Alto Rendimiento (52.3%,  $n=251$ ) **no difiere significativamente** de la proporción esperada del 50% ( $p = 0.338$ ). Aunque existe una ligera diferencia entre los grupos (251 casos en Alto vs 229 casos en Bajo), esta diferencia de aproximadamente 2.3 puntos porcentuales no es estadísticamente significativa y puede atribuirse a la variabilidad muestral aleatoria.

La distribución observada es esencialmente equilibrada entre ambos grupos de rendimiento, lo que indica que no existe un sesgo sistemático en la clasificación de los casos.

## a) Racha

### Datos del Estudio

- **Tamaño de muestra:**  $N = 480$
- **Mediana:** 3.106130
- **Criterio de clasificación:**
  - **0 (Bajo):** Puntajes  $\leq 3.106130$
  - **1 (Alto):** Puntajes  $> 3.106130$

### Hipótesis

- **Hipótesis Nula ( $H_0$ ):** Los puntajes de rendimiento se distribuyen de manera aleatoria en la secuencia de datos. No existe ningún patrón o tendencia sistemática.
- **Hipótesis Alternativa ( $H_1$ ):** Los puntajes de rendimiento NO se distribuyen aleatoriamente. Existe un patrón o tendencia sistemática en la secuencia de datos.

### Nivel de Significación

$$\alpha = 0.05$$

## Pruebas NPar

Estadísticos descriptivos					
	N	Media	Desviación estándar	Mínimo	Máximo
puntaje	480	2992516,56	632700,911	287502	3442203

### Prueba de rachas

	puntaje
Valor de prueba <sup>a</sup>	3091879
Casos < Valor de prueba	240
Casos >= Valor de prueba	240
Casos totales	480
Número de rachas	36
Z	-18,733
Sig. asintótica (bilateral)	,000

a. Mediana

## El ritual de la significancia estadística

1	<p>Plantear Hipótesis</p> <p><math>H_0</math>: La secuencia de puntajes InitialTDI es aleatoria (no existe patrón).</p> <p><math>H_1</math>: La secuencia de puntajes InitialTDI no es aleatoria (existe patrón).</p>
2	<p>Establecer un nivel de significancia</p> <p>Nivel de Significancia (alfa) <math>\alpha = 0.05</math></p>
3	<p>Seleccionar estadístico de prueba (Resaltarla o colocarla en negrita)</p> <p><b>Prueba de Rachas</b></p>
4	<p>Elección de prueba estadística</p> <p>Se utiliza la Prueba de Rachas porque:</p> <ul style="list-style-type: none"><li>• Es una prueba no paramétrica</li><li>• Evalúa si una secuencia de datos es aleatoria</li><li>• No requiere supuestos de normalidad</li><li>• Apropiaada para detectar patrones o tendencias en el orden de los datos</li></ul>
5	<p>Estadísticos clave:</p> <ul style="list-style-type: none"><li>• Número de rachas observadas: 36</li><li>• Número de rachas esperadas (bajo aleatoriedad): <math>\approx 241</math></li><li>• Estadístico Z: -18.733</li><li>• Valor p: 0.000</li></ul> <p>Lectura del p-valor:</p> <p>Dado que el p-valor (0.000) es mucho menor que <math>\alpha = 0.05</math>, se rechaza <math>H_0</math>.</p>
6	<p>Toma de decisiones (dar como respuesta una de las Hipótesis)</p> <p>Se rechaza la hipótesis nula.</p> <p>La secuencia de puntajes de rendimiento NO es aleatoria (existe un patrón sistemático).</p>

A un nivel de significancia del 5%, existe evidencia estadísticamente significativa de un patrón no aleatorio en la secuencia de puntajes de rendimiento ( $Z = -18.733$ ,  $p < 0.001$ ).

### Análisis del patrón detectado:

El número extremadamente bajo de rachas (36 observadas vs ~241 esperadas) indica que los valores tienden a agruparse de manera sistemática:

- Los valores bajos tienden a aparecer consecutivamente
- Los valores altos tienden a aparecer consecutivamente
- Hay muy poca alternancia entre valores altos y bajos

### Implicaciones:

Este patrón sugiere que:

1. Existe una estructura sistemática en el orden de recolección de los datos
2. Los datos pueden estar ordenados o agrupados por algún criterio no aleatorio
3. Podría haber un efecto temporal, secuencial o de agrupamiento en la medición
4. Es posible que los datos estén organizados de menor a mayor o viceversa

### Conclusión:

Se concluye que la secuencia de puntajes de rendimiento NO es aleatoria y presenta un patrón sistemático de agrupamiento. Los valores similares tienden a aparecer consecutivamente en lugar de distribuirse de manera aleatoria a lo largo de la secuencia. Esto es consistente con datos que fueron ordenados o que siguieron algún criterio de agrupación durante su recolección o registro.

## b) Prueba de Mann-Whitney U

### Objetivo y Datos

El objetivo de este análisis es comparar la mediana de los puntajes de rendimiento entre dos grupos independientes clasificados como Alto y Bajo Rendimiento.

- **Variable de Contraste:** Puntaje (Variable numérica continua)
- **Variable de Agrupación:** Grupo (1 = Bajo, 2 = Alto)

---

## Hipótesis

- **H<sub>0</sub>:** La mediana de los puntajes de rendimiento es igual entre los dos grupos (Alto y Bajo).
  - **H<sub>1</sub>:** La mediana de los puntajes de rendimiento es diferente entre los dos grupos.
- 

## Nivel de Significación

$$\alpha = 0.05$$

Estadísticos descriptivos

	N	Media	Desviación estándar	Mínimo	Máximo
puntaje	480	2992516,56	632700,911	287502	3442203
grupo	480	1,48	,500	1	2

### Prueba de Mann-Whitney

Rangos

grupo	N	Rango promedio	Suma de rangos
puntaje alto	251	140,46	35256,00
bajo	229	350,15	80184,00
Total	480		

Estadísticos de prueba<sup>a</sup>

	puntaje
U de Mann-Whitney	3630,000
W de Wilcoxon	35256,000
Z	-16,543
Sig. asintótica (bilateral)	,000

a. Variable de agrupación: grupo

## DISCUSIÓN

### Criterio de decisión:

- Si  $p \leq 0.05$ : Se rechaza H<sub>0</sub>
- Si  $p > 0.05$ : No se rechaza H<sub>0</sub>

Dado que el valor p obtenido (0.000) es significativamente menor que el nivel de significancia de 0.05, se rechaza la hipótesis nula (H<sub>0</sub>).

### Interpretación del estadístico Z:

El valor de  $Z = -16.543$  es extremadamente grande (en valor absoluto), lo que indica una diferencia **muy marcada** entre los dos grupos. El signo negativo indica que el grupo con código menor (Alto = grupo con valores más altos) tiene rangos significativamente menores.

---

## CONCLUSIÓN (Interpretación)

Se concluye que existe una **diferencia estadísticamente muy significativa** en los puntajes de rendimiento entre el grupo de Alto Rendimiento y el grupo de Bajo Rendimiento ( $U = 3630.000$ ,  $Z = -16.543$ ,  $p < 0.001$ ).