
Prueba 2 SOL114

Problema I. (10%)

La prueba **PISA** (Programa para la Evaluación Internacional de Estudiantes) es una evaluación global organizada por la **OCDE** que mide las competencias de los estudiantes de 15 años en tres áreas clave: **lectura**, **matemáticas** y **ciencias**. Esta evaluación se realiza cada tres años y busca proporcionar una visión comparativa del rendimiento educativo entre los países.

La siguiente tabla muestra los resultados en la prueba de **matemáticas** del último PISA 2022 para una submuestra de estudiantes de **Chile** y **Dinamarca**, segmentados por el nivel socioeconómico de los estudiantes. Los datos presentados incluyen los **puntajes promedio** y las **desviaciones estándar** para cada cuartil socioeconómico en ambos países.

Estos resultados se **estimaron** a partir de submuestras de $N_c = 1000$ estudiantes en Chile y $N_d = 1000$ estudiantes en Dinamarca. La tabla muestra que, por ejemplo, según los estadísticos muestras los estudiantes más favorecidos en Chile obtuvieron un puntaje promedio en matemáticas de 470 puntos, con una desviación estándar de 70 puntos.

Cuartil Socioeconómico	Chile: $\hat{\mu}$ (s)	Dinamarca: $\hat{\mu}$ (s)
25% Inferior (Desfavorecidos)	369 (85)	450 (80)
Segundo Cuartil	400 (80)	470 (75)
Tercer Cuartil	430 (75)	490 (70)
25% Superior (Favorecidos)	470 (70)	510 (65)

Nota #1: Recuerda que cada cuartil representa $1/4$ de la muestra de cada país.

Nota #2: Estudios científicos indican que el puntaje de un estudiante en la prueba PISA puede ser tratado como una manifestación una variable aleatoria que sigue una distribución *Lomax* (también conocida como Pareto tipo-II).

I.1. Siguiendo el Teorema del Límite Central, caracteriza la distribución muestral del promedio de puntaje para el segundo cuartil socioeconómico en Chile. Dado que no conocemos los parámetros poblacionales usa su estimado muestral.

Respuesta:

I.2. Calcula un intervalo al 95% de confianza para el promedio puntaje promedio del cuartil socioeconómico más rico en Chile y otro para el puntaje promedio del cuartil socioeconómico más pobre en Dinamarca. Explica que significa cada intervalo de confianza.

Respuesta:

I.3. Calcula la probabilidad de que el promedio muestral del puntaje el cuartil socioeconómico más rico en Chile sobrestime el verdadero promedio en 20 puntos o más.

Respuesta:

I.4. Un político chileno del partido UTI afirma: “Nuestros estudiantes más aventajados (socioeconómicamente) tienen un desempeño en matemáticas superior al de los estudiantes menos aventajados de países socialistas como Dinamarca”. Desarrolla un test de hipótesis para testear esta afirmación con un nivel de significación del 5%.

Respuesta:

I.5. Teniendo en mente la afirmación del político chileno, compara los resultados obtenidos en I.2 e I.4. Explica las potenciales diferencias y similitudes entre ambos.

Respuesta:

Problema II.

Un estudio epidemiológico fue llevado a cabo para evaluar la efectividad de una vacuna contra una enfermedad infecciosa en una comunidad de la Región Metropolitana. Los investigadores hicieron seguimiento durante un año a un grupo de personas, de las cuales algunas recibieron la vacuna como medida preventiva (grupo vacunado), mientras que otras no la recibieron (grupo de control). Los resultados muestran que en el grupo de control, **7 de 10 personas** contrajeron la enfermedad, mientras que **3 no la contrajeron**. Por otro lado, en el grupo vacunado, solo **4 de 10 personas** contrajeron la enfermedad, mientras que **6 no la contrajeron**.

La siguiente tabla resume los resultados

	Contrajo la enfermedad	No contrajo la enfermedad
Grupo Control	7	3
Grupo Vacunado	4	6

Aunque los resultados son preliminares y basados en una muestra pequeña, sugieren un posible beneficio de la vacunación en la reducción de la incidencia de la enfermedad en esta comunidad.

II.1 Testea la hipótesis de existe un beneficio de la vacunación en la reducción de la incidencia de la enfermedad. Trabaja al 95% de confianza.

Respuesta:

II.2 En este estudio se midió además la gravedad de los síntomas en una escala continua (de 0 a 100, donde 0 indica ausencia total de síntomas y 100 los síntomas más graves observados). Los resultados mostraron que el grupo vacunado presentó una media de severidad de síntomas de 30, mientras que el grupo de control presentó una media de 70, sugiriendo una posible reducción en la gravedad de la enfermedad en los vacunados.

En base a estos datos, testea la hipótesis de que existe un beneficio de la vacunación en la reducción de la severidad de los síntomas. Trabaja al 95% de confianza.

Importante: Considera 16 grados de libertad para la diferencia de medias ($df = 16$).

La siguiente tabla resume los resultados:

	Media de severidad de síntomas	$\sum(x_i - \bar{x})^2$	n
Grupo Control	70	1800	10
Grupo Vacunado	30	900	10

Respuesta:

II.3. Un laboratorio desea repetir este experimento, garantizando que, con un nivel de confianza del 95%, el intervalo de confianza para la proporción de personas en cada uno de los grupos tenga un margen de error de, como máximo, ± 0.01 (1 punto porcentual). Suponiendo que tanto el grupo de tratamiento como el de control tendrán el mismo tamaño, calcula el tamaño de muestra necesario para cada grupo para lograr este margen de error.

Respuesta:

Probabilidad acumulada en distribución Normal Standard (mitad positiva)

$Z=z$	$\Phi(z)$
0.00	0.50000
0.01	0.50399
0.02	0.50798
0.03	0.51197
0.04	0.51595
0.05	0.51994
0.06	0.52392
0.07	0.52790
0.08	0.53188
0.09	0.53586
0.10	0.53983
0.11	0.54380
0.12	0.54776
0.13	0.55172
0.14	0.55567
0.15	0.55962
0.16	0.56356
0.17	0.56749
0.18	0.57142
0.19	0.57535
0.20	0.57926
0.21	0.58317
0.22	0.58706
0.23	0.59095
0.24	0.59483
0.25	0.59871
0.26	0.60257
0.27	0.60642
0.28	0.61026
0.29	0.61409
0.30	0.61791
0.31	0.62172
0.32	0.62552

$Z=z$	$\Phi(z)$
0.33	0.62930
0.34	0.63307
0.35	0.63683
0.36	0.64058
0.37	0.64431
0.38	0.64803
0.39	0.65173
0.40	0.65542
0.41	0.65910
0.42	0.66276
0.43	0.66640
0.44	0.67003
0.45	0.67364
0.46	0.67724
0.47	0.68082
0.48	0.68439
0.49	0.68793
0.50	0.69146
0.51	0.69497
0.52	0.69847
0.53	0.70194
0.54	0.70540
0.55	0.70884
0.56	0.71226
0.57	0.71566
0.58	0.71904
0.59	0.72240
0.60	0.72575
0.61	0.72907
0.62	0.73237
0.63	0.73565
0.64	0.73891
0.65	0.74215
0.66	0.74537
0.67	0.74857
0.68	0.75175
0.69	0.75490
0.70	0.75804
0.71	0.76115

$Z=z$	$\Phi(z)$
0.72	0.76424
0.73	0.76730
0.74	0.77035
0.75	0.77337
0.76	0.77637
0.77	0.77935
0.78	0.78230
0.79	0.78524
0.80	0.78814
0.81	0.79103
0.82	0.79389
0.83	0.79673
0.84	0.79955
0.85	0.80234
0.86	0.80511
0.87	0.80785
0.88	0.81057
0.89	0.81327
0.90	0.81594
0.91	0.81859
0.92	0.82121
0.93	0.82381
0.94	0.82639
0.95	0.82894
0.96	0.83147
0.97	0.83398
0.98	0.83646
0.99	0.83891
1.00	0.84134
1.01	0.84375
1.02	0.84614
1.03	0.84849
1.04	0.85083
1.05	0.85314
1.06	0.85543
1.07	0.85769
1.08	0.85993
1.09	0.86214
1.10	0.86433

$Z=z$	$\Phi(z)$
1.11	0.86650
1.12	0.86864
1.13	0.87076
1.14	0.87286
1.15	0.87493
1.16	0.87698
1.17	0.87900
1.18	0.88100
1.19	0.88298
1.20	0.88493
1.21	0.88686
1.22	0.88877
1.23	0.89065
1.24	0.89251
1.25	0.89435
1.26	0.89617
1.27	0.89796
1.28	0.89973
1.29	0.90147
1.30	0.90320
1.31	0.90490
1.32	0.90658
1.33	0.90824
1.34	0.90988
1.35	0.91149
1.36	0.91309
1.37	0.91466
1.38	0.91621
1.39	0.91774
1.40	0.91924
1.41	0.92073
1.42	0.92220
1.43	0.92364
1.44	0.92507
1.45	0.92647
1.46	0.92785
1.47	0.92922
1.48	0.93056
1.49	0.93189

$Z=z$	$\Phi(z)$
1.50	0.93319
1.51	0.93448
1.52	0.93574
1.53	0.93699
1.54	0.93822
1.55	0.93943
1.56	0.94062
1.57	0.94179
1.58	0.94295
1.59	0.94408
1.60	0.94520
1.61	0.94630
1.62	0.94738
1.63	0.94845
1.64	0.94950
1.65	0.95053
1.66	0.95154
1.67	0.95254
1.68	0.95352
1.69	0.95449
1.70	0.95543
1.71	0.95637
1.72	0.95728
1.73	0.95818
1.74	0.95907
1.75	0.95994
1.76	0.96080
1.77	0.96164
1.78	0.96246
1.79	0.96327
1.80	0.96407
1.81	0.96485
1.82	0.96562
1.83	0.96638
1.84	0.96712
1.85	0.96784
1.86	0.96856
1.87	0.96926
1.88	0.96995

$Z=z$	$\Phi(z)$
1.89	0.97062
1.90	0.97128
1.91	0.97193
1.92	0.97257
1.93	0.97320
1.94	0.97381
1.95	0.97441
1.96	0.97500
1.97	0.97558
1.98	0.97615
1.99	0.97670
2.00	0.97725
2.01	0.97778
2.02	0.97831
2.03	0.97882
2.04	0.97932
2.05	0.97982
2.06	0.98030
2.07	0.98077
2.08	0.98124
2.09	0.98169
2.10	0.98214
2.11	0.98257
2.12	0.98300
2.13	0.98341
2.14	0.98382
2.15	0.98422
2.16	0.98461
2.17	0.98500
2.18	0.98537
2.19	0.98574
2.20	0.98610
2.21	0.98645
2.22	0.98679
2.23	0.98713
2.24	0.98745
2.25	0.98778
2.26	0.98809
2.27	0.98840

$Z=z$	$\Phi(z)$
2.28	0.98870
2.29	0.98899
2.30	0.98928
2.31	0.98956
2.32	0.98983
2.33	0.99010
2.34	0.99036
2.35	0.99061
2.36	0.99086
2.37	0.99111
2.38	0.99134
2.39	0.99158
2.40	0.99180
2.41	0.99202
2.42	0.99224
2.43	0.99245
2.44	0.99266
2.45	0.99286
2.46	0.99305
2.47	0.99324
2.48	0.99343
2.49	0.99361
2.50	0.99379
2.51	0.99396
2.52	0.99413
2.53	0.99430
2.54	0.99446
2.55	0.99461
2.56	0.99477
2.57	0.99492
2.58	0.99506
2.59	0.99520
2.60	0.99534
2.61	0.99547
2.62	0.99560
2.63	0.99573
2.64	0.99585
2.65	0.99598
2.66	0.99609

$Z=z$	$\Phi(z)$
2.67	0.99621
2.68	0.99632
2.69	0.99643
2.70	0.99653
2.71	0.99664
2.72	0.99674
2.73	0.99683
2.74	0.99693
2.75	0.99702
2.76	0.99711
2.77	0.99720
2.78	0.99728
2.79	0.99736
2.80	0.99744
2.81	0.99752
2.82	0.99760
2.83	0.99767
2.84	0.99774
2.85	0.99781
2.86	0.99788
2.87	0.99795
2.88	0.99801
2.89	0.99807
2.90	0.99813
2.91	0.99819
2.92	0.99825
2.93	0.99831
2.94	0.99836
2.95	0.99841
2.96	0.99846
2.97	0.99851
2.98	0.99856
2.99	0.99861
3.00	0.99865

Probabilidad acumulada en distribución t Standard (mitad positiva)

t_value	CDF for df = 1
0.0	0.50000
0.1	0.53173
0.2	0.56283
0.3	0.59277
0.4	0.62112
0.5	0.64758
0.6	0.67202
0.7	0.69440
0.8	0.71478
0.9	0.73326
1.0	0.75000
1.1	0.76515
1.2	0.77886
1.3	0.79129
1.4	0.80257
1.5	0.81283
1.6	0.82219
1.7	0.83075
1.8	0.83859
1.9	0.84579
2.0	0.85242
2.1	0.85854
2.2	0.86420
2.3	0.86945
2.4	0.87433
2.5	0.87888
2.6	0.88312
2.7	0.88709
2.8	0.89081
2.9	0.89430
3.0	0.89758

t_value	CDF for df = 5
:-----:	:-----:
0.0	0.50000
0.1	0.53788
0.2	0.57532
0.3	0.61188
0.4	0.64716
0.5	0.68085
0.6	0.71267
0.7	0.74243
0.8	0.76999
0.9	0.79531
1.0	0.81839
1.1	0.83927
1.2	0.85805
1.3	0.87485
1.4	0.88980
1.5	0.90305
1.6	0.91475
1.7	0.92506
1.8	0.93412
1.9	0.94207
2.0	0.94903
2.1	0.95512
2.2	0.96045
2.3	0.96511
2.4	0.96919
2.5	0.97275
2.6	0.97588
2.7	0.97861
2.8	0.98100
2.9	0.98310
3.0	0.98495

t_value	CDF for df = 10
:-----:	:-----:
0.0	0.50000
0.1	0.53884

	0.2		0.57726	
	0.3		0.61484	
	0.4		0.65122	
	0.5		0.68605	
	0.6		0.71907	
	0.7		0.75006	
	0.8		0.77885	
	0.9		0.80536	
	1.0		0.82955	
	1.1		0.85145	
	1.2		0.87110	
	1.3		0.88862	
	1.4		0.90412	
	1.5		0.91775	
	1.6		0.92966	
	1.7		0.94002	
	1.8		0.94897	
	1.9		0.95669	
	2.0		0.96331	
	2.1		0.96896	
	2.2		0.97378	
	2.3		0.97787	
	2.4		0.98134	
	2.5		0.98428	
	2.6		0.98675	
	2.7		0.98884	
	2.8		0.99060	
	2.9		0.99208	
	3.0		0.99333	

	t_value		CDF for df = 15	
	:-----:		:-----:	
	0.0		0.50000	
	0.1		0.53917	
	0.2		0.57792	
	0.3		0.61585	
	0.4		0.65260	
	0.5		0.68783	

	0.6		0.72127	
	0.7		0.75268	
	0.8		0.78190	
	0.9		0.80883	
	1.0		0.83341	
	1.1		0.85566	
	1.2		0.87563	
	1.3		0.89339	
	1.4		0.90907	
	1.5		0.92282	
	1.6		0.93478	
	1.7		0.94512	
	1.8		0.95400	
	1.9		0.96158	
	2.0		0.96803	
	2.1		0.97347	
	2.2		0.97805	
	2.3		0.98189	
	2.4		0.98509	
	2.5		0.98775	
	2.6		0.98995	
	2.7		0.99177	
	2.8		0.99327	
	2.9		0.99450	
	3.0		0.99551	

	t_value		CDF for df = 16	
	:-----:		:-----:	
	0.0		0.50000	
	0.1		0.53921	
	0.2		0.57800	
	0.3		0.61598	
	0.4		0.65278	
	0.5		0.68806	
	0.6		0.72155	
	0.7		0.75301	
	0.8		0.78229	
	0.9		0.80927	

	1.0		0.83390	
	1.1		0.85620	
	1.2		0.87620	
	1.3		0.89399	
	1.4		0.90970	
	1.5		0.92346	
	1.6		0.93542	
	1.7		0.94576	
	1.8		0.95463	
	1.9		0.96220	
	2.0		0.96861	
	2.1		0.97403	
	2.2		0.97858	
	2.3		0.98238	
	2.4		0.98554	
	2.5		0.98816	
	2.6		0.99033	
	2.7		0.99211	
	2.8		0.99358	
	2.9		0.99478	
	3.0		0.99576	