

Prueba de Kolmogorov-Smirnov. D

Kolmogorov-Smirnov supusieron que la distribución de probabilidad que se encontraba a prueba, era continua y que se conocía la media y la varianza de la población.

La prueba se emplea para probar el grado de concordancia entre la distribución de datos empíricos de la muestra y alguna distribución teórica específica.

Pasos de la prueba.

1. Se desarrolla la distribución acumulativa $F(x)$ de la distribución teórica y de los datos empíricos tales que $n < 100$.
2. Se calculan las desviaciones D y se selecciona la mayor desviación absoluta (desviación máxima) D_{\max} .
3. Se compara la desviación máxima D_{\max} obtenida con la desviación crítica D_{teo} obtenida de la tabla de Kolmogorov-Smirnov.

Si la desviación observada es **menor** a la desviación **crítica** se acepta que los datos observados no presentan ninguna diferencia significativa con los que proporciona la distribución teórica con ciertos parámetros.

ejemplo. Se obtuvo una muestra de las ventas, cuyo comportamiento se considera **Normal**.

clase	x	f _i obs	f(x) obs	F(x) obs acum	Z lim inf	P lim inf	P teo normal	P teo acum	D
20-30	25	5	.0632	0.0632	-2.4182	.0080	.0502	.0502	.0130
30-40	35	12	.1518	0.2151	-1.5785	.0582	.1745	.2247	.0096
40-50	45	26	.3291	0.5443	-.7387	.2327	.3071	.5318	.0125
50-60	55	23	.2911	0.8354	.1009	.4602	.2866	.8184	.0170
60-70	65	10	.1265	0.9620	.9407	.1736	.1361	.9545	.0075
70-80	75	3	.0379	1	1.7805	.0375	.0331	.9876	.0124
		79	1.000		2.6202	.0044			

Media=3855/79=48.7975 s=11.9082

$$Z = (x - \text{media}) / s$$

$$Z_{20} = (20 - 48.7975) / 11.9082 = -2.4182$$

$$Z_{30} = (30 - 48.7975) / 11.9082 = -1.5785 \dots$$

$$Z_{80} = (80 - 48.7975) / 11.9082 = 2.6202$$

P teórica de normal

$$P_{25} = 0.0582 - 0.0080 =$$

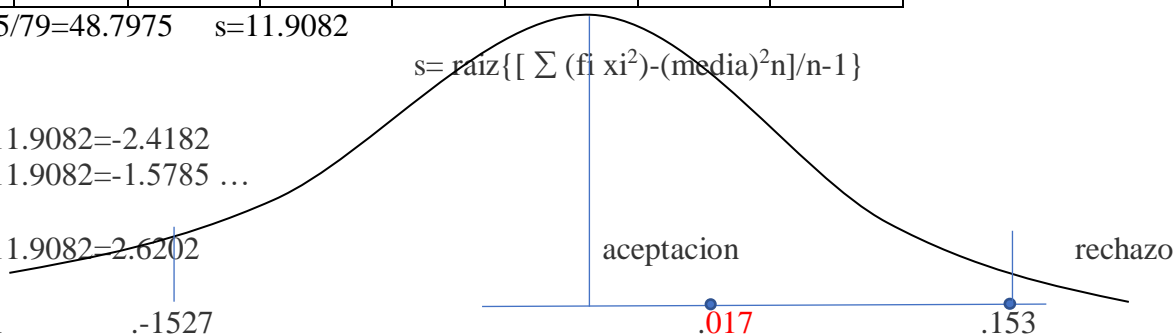
$$P_{35} = 0.2327 - 0.0582 =$$

La mayor **desviación absoluta observada (D max)** es de **0.017**

Buscando en la tabla de Kolmogorov-Smirnov la desviación crítica con $n = 79$ y $\alpha = 0.05$, nos indica que se requiere calcular con la relación: $1.36 / \sqrt{N}$

$$\text{Desviación crítica} = 1.36 / \sqrt{79} = 0.153$$

Como la **desviación absoluta observada (0.017)** es menor que la **desviación crítica tabulada (0.153)**, se acepta la hipótesis nula (la muestra tiene comportamiento Normal).



De la siguiente muestra, determine su comportamiento con una confiabilidad del 95%.

Clase i	marca	Frec obs	Prob obs (i)	Prob obs acum	Z lim inf	Prob Z	Prob teo normal (i)	Prob teo acum	D
10-30	20	5	.0746	.0746	-2.5282	0.0059	0.0571	0.0571	0.0175
30-50	40	12	.1791	.2537	-1.5347	0.063	0.2316	0.2887	0.035
50-70	60	30	.4477	.7014	-0.5412	0.2946	0.379	0.6677	0.0337
70-90	80	15	.2238	.9252	0.4522	0.3264	0.2545	0.9222	0.003
90-110	100	5	.0746	.9998	1.4457	0.0749	0.0674	0.9896	0.0102

$$Z_{110} = 2.4392 \quad 0.0075$$

$$n=67 \quad x=4080/67=60.8955 \quad s^2 = [275 \ 200 - 248 \ 453.5486]/(n-1)=405.2492 \quad s= 20.1308$$

$$Z_{10}=(10-60.8955)/20.1308=-2.5282$$

$$Z_{30}=(30-60.8955)/20.1308=-1.5347$$

$$Z_{50}=(50-60.8955)/20.1308=-0.5412$$

$$|D|_{\max} = 0.035$$

$$|D|_{\text{teo}}(n, \alpha) = 0.1991$$

EL COMPORTAMIENTO DE LA MUESTRA SI ES NORMAL

De la siguiente muestra, determine su comportamiento con una confiabilidad del 95%. Usando la distribución poisson.

Clase i	Frec obs	Prob obs (i)	Pro acum obs	Prob teo poisson(i)	Prob teo acum	D
0	60	0.6521	.6521	0.5540	.554	.0981
1	20	0.2173	.8694	0.3271	.8811	.0117
2	10	0.1086	.9722	0.0965	.9776	.0054
3	1	0.0108	.983	0.019	.9966	.0136
4	1	0.0108	.9938	0.0028	.9994	.0056

$$n=92 \quad x=47/92=0.5108$$

$$s^2 = (85 - 24.0043)/91=0.6702$$

$$\lambda = (.5108+.6702)/ 2=0.5905$$

$$P(0) = [(0.5905) (2.71828)^{-(0.5905*0)}] / 0! = 0.5540$$

$$P(1) = [(0.5905)^1 (2.71828)^{-(0.5905)}] / 1! = 0.3271$$

$$P(2) = 0.0965$$

$$P(3) = 0.0190$$

$$P(4) = 0.0028$$

$$|D|_{\max}=0.0981$$

$$|D|_{\text{teo}}(n, \alpha)=0.1417$$

la muestra tiene comportamiento Poisson

De la siguiente muestra, determine su comportamiento con una confiabilidad del 95%. Usando la distribución exponencial.

Clase i	Frec obs	Prob obs (i)	Pro acum obs	Prob teo Expon(i)	Prob teo acum	D
0	30	.6	.6	.6	.6	0
1	13	.26	.86	.3292	.9292	.0692
2	05	.1	.96	.1807	1.1099	.1499
3	1	.02	.98	.0991	1.209	.229
4	1	.02	1	.0544	1.2634	.2634

$$n=50 \quad x=30/50=0.6 \quad s^2=(58-(0.6^2(50))/49=0.8163 \quad \lambda=0.6$$

$$P(0) = [(0.6) (2.71828)^{-(0.6(0))}] = 0.6$$

$$P(1) = [(0.6) (2.71828)^{-(0.6(1))}] = 0.3292$$

$$P(2) = [(0.6) (2.71828)^{-(0.6(2))}] = 0.1807$$

$$P(3) = [(0.6) (2.71828)^{-(0.6(3))}] = 0.0991$$

$$P(4) = [(0.6) (2.71828)^{-(0.6(4))}] = 0.0544$$

$$|D|_{\max}=0.2634$$

$$|D|_{\text{teo}}(n, \alpha)=0.1923$$

la muestra NO tiene comportamiento EXP

De la siguiente muestra, determine su comportamiento con una confiabilidad del 95%. Usando la distribución exponencial.

Clase i	Frec obs	Prob obs (i)	Pro acum obs	Prob teo exponencial(i)	Prob teo acum	D
0	60	0.6521	.6521	0.5108	.5108	.1413
1	20	0.2173	.8694	0.3064	.8174	.0522
2	10	0.1086	.9722	0.1838	1.003	.0308
3	1	0.0108	.983	0.1103	1.1133	.1303
4	1	0.0108	.9938	0.0662	1.1678	.174

$$n=92 \quad x=47/92=0.5108 \quad s^2=(85-24.0043)/91=0.6702 \quad \lambda=0.5108$$

$$P(0) = [(0.5108) (2.71828)^{-(0.5108)0}] = 0.5108$$

$$P(1) = [(0.5108) (2.71828)^{-(0.5108)1}] = 0.3064$$

$$P(2) = [(0.5108) (2.71828)^{-(0.5108)2}] = 0.1838$$

$$P(3) = [(0.5108) (2.71828)^{-(0.5108)3}] = 0.1103$$

$$P(4) = [(0.5108) (2.71828)^{-(0.5108)4}] = 0.0662$$

$$|D|_{\max}=0.174$$

$$|D|_{\text{teo}}(n, \alpha)=0.1417$$

la muestra NO tiene comportamiento EXP