



Prüfung der
theorie
abhängig.

Vale la pena destacar que en todos los problemas vistos hasta ahora, había un solo un factor que captaba el interés. Sin embargo, chi-cuadrado también permitirá la comparación de dos atributos para determinar si existe una relación entre ellos.

Ejemplo 1

Un minorista desea descubrir si existe una relación entre los niveles de ingreso de los consumidores y su preferencia por su producto. Los gerentes de producción siempre se interesan por la relación que pueda existir entre la productividad de sus empleados y el tipo o grado de capacitación que recibieron. La directora de investigación de productos en Dow Chemical. En su proyecto actual, la debe determinar si existe alguna relación entre la clasificación de efectividad que los consumidores asignan a un nuevo insecticida y el sitio (urbano o rural) en el cual se utiliza. La tabla resume las clasificaciones hechas por cada consumidor en una tabla de contingencia como las que se presentaron con anterioridad.

	Atributo B -Ubicación		
Atributo A -Clasificación	Urbano	Rural	Total
por encima del promedio	20	11	31
promedio	40	8	48
por debajo del promedio	15	6	21
Total	75	25	100

La tabla tiene $f = 3$ filas y $c = 2$ columnas. Existen $f \times c = 6$ celdas en la tabla. Por ejemplo, vale la pena observar que 31 clientes clasificaron el producto "por encima del promedio", 20 de ellos estaban en zonas urbanas.

La directora desea comparar el atributo B (ubicación) con el atributo A (clasificación del producto). Sus hipótesis son:

Ensayo de hipótesis

H_0 ; La clasificación y la ubicación son independientes

H_1 ; La clasificación y la ubicación no son independientes.

Si la ubicación no tiene ningún impacto en la clasificación de efectividad, entonces el porcentaje de residentes **urbanos** que calificaron el producto "por encima del promedio" debería ser igual al porcentaje de residentes **rurales** que clasificaron el producto "por encima del promedio". Este porcentaje a su vez debería ser igual al de **todos** los usuarios que calificaron el producto "por encima del promedio".

Como lo muestra la tabla, el 31% de todos los usuarios clasificaron el producto "por encima del promedio". Luego el 31% de los 75 residentes **urbanos** y el 31% de los 25 residentes **rurales** deberán dar esta clasificación si la clasificación y la ubicación son independientes.

Estos valores de $(75)(0.31) = 23.3$ y $(25)(0.31) = 7.75$

dan las frecuencias esperadas E_i para cada celda, como se observa en la siguiente tabla

Atributo A -Clasificación	Atributo B -Ubicación		Total
	Urbano	Rural	
por encima del promedio	$O_i=20$ $E_i=23.25$	$O_i=11$ $E_i=7.75$	31
promedio	$O_i=40$ $E_i=36$	$O_i=8$ $E_i=12$	48
por debajo del promedio	$O_i=15$ $E_i=15.75$	$O_i=6$ $E_i=5.25$	21
Total	75	25	100

Cálculo Probar las hipótesis requiere una comparación de O_i y E_i , sobre las $fc = 6$ celdas, utilizamos la ecuación:

$$\chi^2 = \sum_{i=1}^{fc} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = \sum_{i=1}^6 \frac{(O_i - E_i)^2}{E_i} = \frac{(20 - 23.25)^2}{23.25} + \frac{(40 - 36)^2}{36} + \frac{(15 - 15.8)^2}{15.8} + \frac{(11 - 7.75)^2}{7.75} + \frac{(8 - 12)^2}{12} + \frac{(6 - 5.25)^2}{5.25} = 3.76$$

La prueba tiene $(f-1) \times (c-1) = (3 - 1)(2 - 1) = 2$ grados de libertad. Si la directora fija $\alpha = 10\%$, tenemos $\chi^2_{0.10,2} = 4.605$.

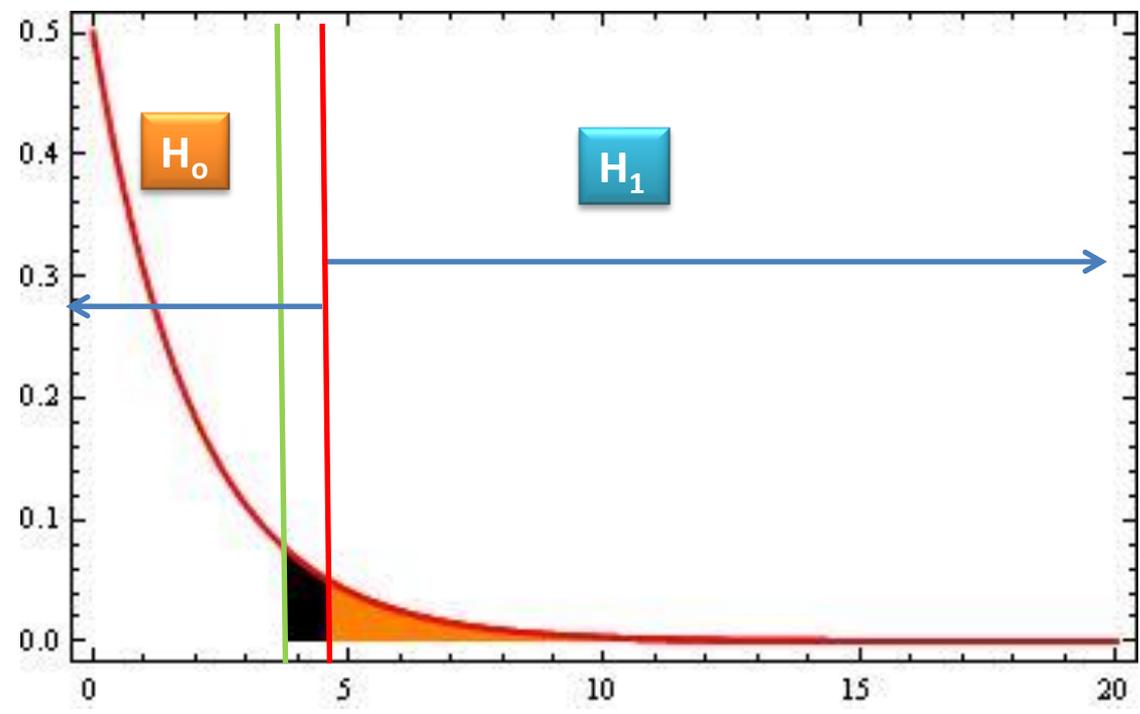
g.l.	$\chi^2_{0.995}$	$\chi^2_{0.900}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$	$\chi^2_{0.700}$	$\chi^2_{0.500}$	$\chi^2_{0.300}$	$\chi^2_{0.200}$	$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.020}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$
1	0.000	0.000	0.001	0.004	0.016	0.148	0.455	1.074	1.642	2.706	3.841	5.024	5.412	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.713	1.386	2.408	3.219	4.605	5.991	7.378	7.824	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	1.424	2.366	3.665	4.642	6.251	7.815	9.348	9.837	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	2.195	3.357	4.878	5.989	7.779	9.488	11.143	11.668	13.277	14.860

Regla de decisión:

No rechazar H_0 si $\chi^2 < 4.605$.
Rechazar H_0 si $\chi^2 > 4.605$.

Conclusión y justificación

El valor obtenido de $\chi^2 = 3.76$, nos permite concluir con un nivel de significancia de 10% que no existe relación entre clasificación y ubicación, es decir ambos atributos son independientes.



Ejemplo 2

“Seat Autoemoción” desea determinar si existe alguna relación entre el ingreso de las clientes y la importancia que dan al precio de los automóviles de lujo. La gerente de la compañía desea probar la hipótesis:

H_0 : Ingreso e importancia del precio son independientes.

H_1 : Ingreso e importancia del precio no son independientes.

Los clientes están agrupados en tres niveles de ingreso y se les pide asignar un nivel de significancia para poner precio a la decisión de compra. Los resultados se muestran en la siguiente tabla de contingencia.

Atributo A - Nivel de Importancia	Atributo B - Ingreso			Total
	Bajo	Medio	Alto	
Grande	$O_i=83$	$O_i=62$	$O_i=37$	182
	$E_i=66.98$	$E_i=$		
Moderado	$O_i=52$	$O_i=71$	$O_i=49$	172
	$E_i=$	$E_i=$		
Poco	$O_i=63$	$O_i=58$	$O_i=63$	184
	$E_i=$	$E_i=$		
Total	198	191	149	538

Debido a que $182/538 = 33.83\%$ de todos las que respondieron a la encuesta agregan un nivel de importancia "grande" al precio, entonces si el ingreso y el precio no es tan relacionados, se esperaría que 33.83% de ellos, en cada clasificación de ingresos respondan que el precio era de "gran" importancia. Por tanto, los E_i para un nivel de importancia "bajo" son:

$$(198) \times ((82/538) = 0.3383) = 66.98,$$

$$(191) \times (0.3383) = 64.62,$$

$$(149) \times (0.3383) = 50.41$$

De forma similar, $172/538 = 31.97\%$ de todos quienes respondieron clasificar el precio como moderadamente importante. Por ende las E_i para la categoría de "moderado" son

$$198 \times (0.3197) = 63.32,$$

$$191 \times (0.3197) = 61.06,$$

$$149 \times (0.3197) = 47.64.$$

Para la categoría de "poco", los datos muestran que $184/538 = 34.20\%$ de todas las clientes daban poca importancia al precio. Por consiguiente los E_i son:

$$198 \times (0.342) = 67.72,$$

$$191 \times (0.342) = 65.32,$$

$$149 \times (0.342) = 50.96.$$

Ensayo de hipótesis

H_0 ; Ingreso e importancia del precio son independientes.

H_1 ; Ingreso e importancia del precio no son independientes

Chi-cuadrado es

Cálculo

$$\chi^2 = \sum_{i=1}^9 \frac{(O_i - E_i)^2}{E_i} = \frac{(83 - 66.98)^2}{66.98} + \frac{(52 - 63.32)^2}{63.32} + \frac{(63 - 67.72)^2}{67.72} + \frac{(62 - 64.62)^2}{64.62} + \frac{(71 - 61.06)^2}{61.06} + \frac{(49 - 47.64)^2}{47.64} + \frac{(37 - 50.41)^2}{50.41} + \frac{(49 - 47.64)^2}{47.64} + \frac{(63 - 50.96)^2}{50.96} = 15.17$$

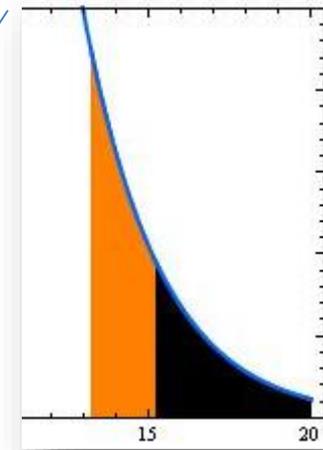
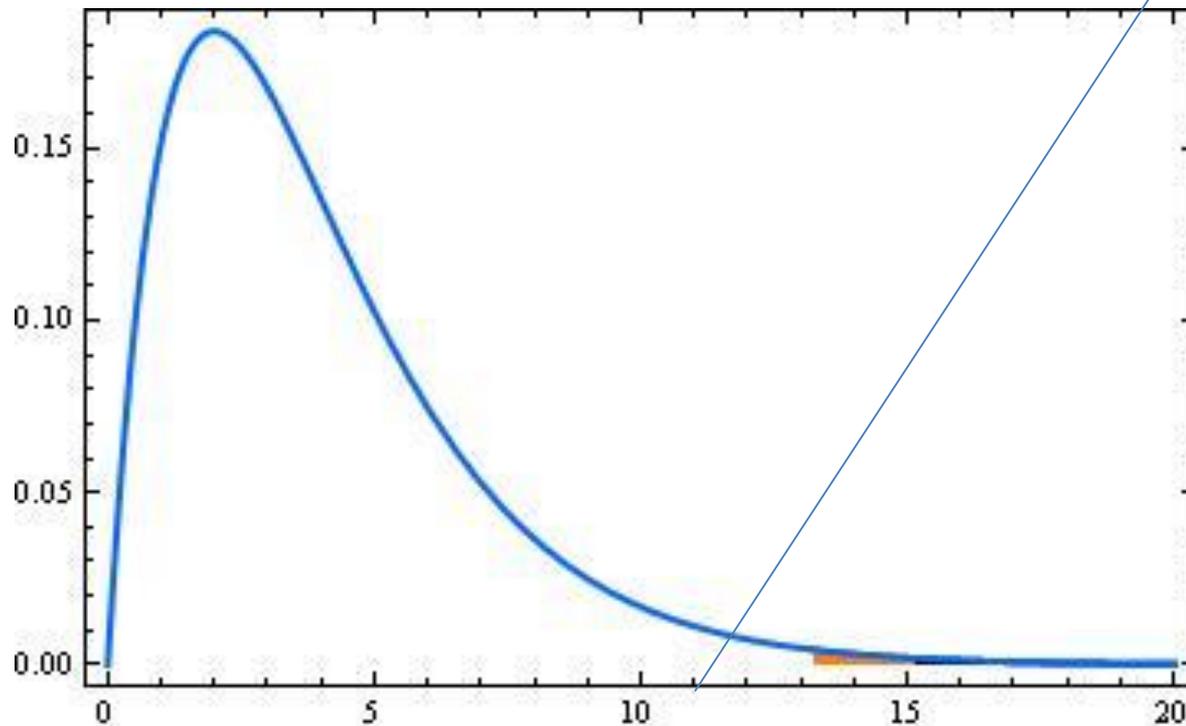
Si Se determina α en 1%, y con $gl = (f - 1) \times (c - 1) = (3 - 1) \times (3 - 1) = 4$ gl. Tenemos que buscar en la tabla para la función chi cuadrado $\chi^2_{0.01,4}$

g.l.	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$	$\chi^2_{0.700}$	$\chi^2_{0.500}$	$\chi^2_{0.300}$	$\chi^2_{0.200}$	$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.020}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$
1	0.000	0.000	0.001	0.004	0.016	0.148	0.455	1.074	1.642	2.706	3.841	5.024	5.412	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.713	1.386	2.408	3.219	4.605	5.991	7.378	7.824	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	1.424	2.366	3.665	4.642	6.251	7.815	9.348	9.837	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	2.195	3.357	4.878	5.989	7.779	9.488	11.143	11.668	13.277	14.860

Como se observa en la figura, la regla de decisión es:

Regla de decisión:

No rechazar H_0 si $\chi^2 \leq 13.277$
Rechazar H_0 si $\chi^2 > 13.277$



La hipótesis nula se rechaza. Existe sólo 1% de probabilidad de que si no existe relación entre ingreso y significancia del precio, las diferencias entre O_i y E_i serían lo suficiente mente grandes como para producir un chi-cuadrado mas grande que 13.277. Existe evidencia de una relación entre el ingreso de los clientes y la importancia dada al precio de un auto de lujo.

Las próximas clases dedicaremos nuestros esfuerzos a pruebas no paramétricas. Varias de ellas se analizan para ilustrar tres puntos importantes:

- 1) Cuando debe usarse la prueba,
- 2) Su propósito principal,
- 3) La interpretación.

Algunas de estas pruebas no paramétricas son:

- 1) La prueba de signo,
- 2) La prueba de rechas,
- 3) La prueba de U de Mann-Whitney
- 4) La prueba de correlación de rengos de Spearman.
- 5) La prueba de Krustal-Wallis

The End



g.l.	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$	$\chi^2_{0.700}$	$\chi^2_{0.500}$	$\chi^2_{0.300}$	$\chi^2_{0.200}$	$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.020}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$
1	0.000	0.000	0.001	0.004	0.016	0.148	0.455	1.074	1.642	2.706	3.841	5.024	5.412	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.713	1.386	2.408	3.219	4.605	5.991	7.378	7.824	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	1.424	2.366	3.665	4.642	6.251	7.815	9.348	9.837	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	2.195	3.357	4.878	5.989	7.779	9.488	11.143	11.668	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	3.000	4.351	6.064	7.289	9.236	11.070	12.833	13.388	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	3.828	5.348	7.231	8.558	10.645	12.592	14.449	15.033	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	4.671	6.346	8.383	9.803	12.017	14.067	16.013	16.622	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	5.527	7.344	9.524	11.030	13.362	15.507	17.535	18.168	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	6.393	8.343	10.656	12.242	14.684	16.919	19.023	19.679	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	7.267	9.342	11.781	13.442	15.987	18.307	20.483	21.161	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	8.148	10.341	12.899	14.631	17.275	19.675	21.920	22.618	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	9.034	11.340	14.011	15.812	18.549	21.026	23.337	24.054	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	9.926	12.340	15.119	16.985	19.812	22.362	24.736	25.472	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	10.821	13.339	16.222	18.151	21.064	23.685	26.119	26.873	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	11.721	14.339	17.322	19.311	22.307	24.996	27.488	28.259	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	12.624	15.338	18.418	20.465	23.542	26.296	28.845	29.633	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	13.531	16.338	19.511	21.615	24.769	27.587	30.191	30.995	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	14.440	17.338	20.601	22.760	25.989	28.869	31.526	32.346	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	15.352	18.338	21.689	23.900	27.204	30.144	32.852	33.687	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	16.266	19.337	22.775	25.038	28.412	31.410	34.170	35.020	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	17.182	20.337	23.858	26.171	29.615	32.671	35.479	36.343	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	18.101	21.337	24.939	27.301	30.813	33.924	36.781	37.659	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	19.021	22.337	26.018	28.429	32.007	35.172	38.076	38.968	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	19.943	23.337	27.096	29.553	33.196	36.415	39.364	40.270	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	20.867	24.337	28.172	30.675	34.382	37.652	40.646	41.566	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	21.792	25.336	29.246	31.795	35.563	38.885	41.923	42.856	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	22.719	26.336	30.319	32.912	36.741	40.113	43.194	44.140	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	23.647	27.336	31.391	34.027	37.916	41.337	44.461	45.419	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	24.577	28.336	32.461	35.139	39.087	42.557	45.722	46.693	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	25.508	29.336	33.530	36.250	40.256	43.773	46.979	47.962	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	34.872	39.335	44.165	47.269	51.805	55.758	59.342	60.436	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	44.313	49.335	54.723	58.164	63.167	67.505	71.420	72.613	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	53.809	59.335	65.227	68.972	74.397	79.082	83.298	84.580	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	63.346	69.334	75.689	79.715	85.527	90.531	95.023	96.388	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	72.915	79.334	86.120	90.405	96.578	101.879	106.629	108.069	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	82.511	89.334	96.524	101.054	107.565	113.145	118.136	119.648	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	92.129	99.334	106.906	111.667	118.498	124.342	129.561	131.142	135.807	140.169