TC2007 Métodos Cuantitativos y Simulación

Números aleatorios

Los números aleatorios son utilizados en simulación

Para generar eventos

Los lenguajes de programación cuenta con rutinas para la generación de números aleatorios

Propiedades de los números aleatorios

Una secuencia de números aleatorios $R_1, R_2, \ldots, R_n$ deben tener 2 propiedades estadísticas:

- Uniformidad
- Independencia

Generación de números *pseudo*-Aleatorios

Problemas que se presentan al generar números aleatorios

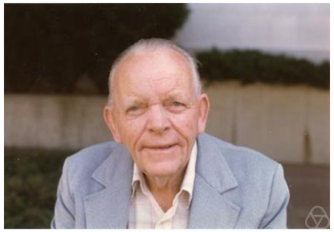
1. Los números generados pueden no estar distribuidos **uniformemente**
2. Los números generados pueden pueden tener valores discretos en lugar de continuos
3. La media de los números generados puede ser muy pequeña o muy grande
4. La varianza de los números generados puede ser muy grande o muy pequeña
5. Puede existir una **dependencia** como autocorrelación

Generación de números *pseudo*-Aleatorios

Los métodos para generar números aleatorios:

1. Deben ser rápidos
2. Deben ser portables en diferentes computadoras
3. Deben tener un ciclo suficientemente grande
4. Debe ser repetible bajo las mismas condiciones, independiente del sistema utilizado
5. Deben tener una aproximación muy cercana a las propiedades estadísticas ideales que son ***uniformidad*** e ***independencia***

Método de Congruencia Lineal



Fue propuesto por *Lehmer* (1951) en la Universidad de Berkeley California

Produce una secuencia de enteros $X_1, X_2, \ldots$ entre ***0*** y ***m-1***, utilizando la expresión recursiva:

$$X_{i+1} = (aX_i + c) \bmod m, \qquad \text{para } i=0,1,2,\ldots$$

X_0 es la semilla

a es el multiplicador $0 \leq a < m$

c es el incremento $0 \leq a < m$

m es el modulo

*Si c=0 el método recibe el nombre de **multiplicativo secuencial***

Se generan Enteros Aleatorios

Los enteros generados tienen una distribución uniforme entre ***0*** y ***m***

Para generar números aleatorio entre ***0*** y ***1***

Para generar números aleatorio entre ***0*** y ***1*** $$R_i = \frac{X_i}{m}, \quad i = 1, 2, \ldots$$

Ejemplo Método de Congruencia Lineal

$X_{i+1} = (aX_i + c) \bmod m,$ para $i=0,1,2,\ldots$

$\boldsymbol{X_0} = 27$ semilla

$\boldsymbol{a} = 17$

$\boldsymbol{c} = 43$

$\boldsymbol{m} = 100$

$X_{i+1} = [(17)(27) + 43] \bmod 100$ $X_{i+1} = 2$ $R_1 = 2/100 = 0.02$

$X_{i+1} = [(17)(2) + 43] \bmod 100$ $X_{i+1} = 77$ $R_2 = 77/100 = 0.77$

$X_{i+1} = [(17)(77) + 43] \bmod 100$ $X_{i+1} = 52$ $R_3 = 52/100 = 0.52$

$X_0 =$ 1, 2, 3, 4
$a = 13$
$c = 0$
$m = 64$

i	X_i	X_i	X_i	X_i
0	1	2	3	4
1	13	26	39	52
2	41	18	59	36
3	21	42	63	20
4	17	34	51	4
5	29	58	23	
6	57	50	43	
7	37	10	47	
8	33	2	35	
9	45		7	
10	9		27	
11	53		31	
12	49		19	
13	61		55	
14	25		11	
15	5		15	
16	1		3	

Método de Congruencia Lineal

Un conjunto de valores ampliamente probado por *[Learmonth and Lewis, 1973; Lewis et al., 1969]* es:

Xo = 123457
a = 16807
c = 0
m = 2147483647

Con estos valores el periodo es de *P=m-1*, más de *2,000,000,000*

X_i	Entero Aleatorio	Número Aleatorio
123457	2074941799	0.96622007
2074941799	559872160	0.260710791
559872160	1645535613	0.766262232
1645535613	1222641625	0.569336873
1222641625	1814256879	0.844829194
1814256879	95061600	0.044266507
95061600	2119961479	0.987183992
2119961479	1291390176	0.601350412
1291390176	1924951450	0.896375371
1924951450	817878095	0.380854167
817878095	34318218	0.015980666

. . .

Prueba de *chi-cuadrada (Uniformidad)*

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

La prueba se aplica a una muestra de datos

Los datos se dividen en *n clases o intervalos*

O_i es el número observado en la i_{th} ***clase***
E_i es el número esperado en la i_{th} ***clase*** (distribución uniforme) E_i=N/n, N total de datos
n es el número de clases

Podemos considerar 10 clases, esto derivado de que estamos con variables aleatorias con un rango de valores entre [0,1], las clases pueden ser clases (*1,2,..,10)*

Esta prueba contrasta frecuencias observadas contra frecuencias esperadas de acuerdo a la hipótesis nula

Para una distribución uniforme, tenemos que E_i es:

$$E_i = \frac{N}{n}$$

N conjunto de datos

Ejemplo:
Utiliza la prueba de *chi-cuadrada* con *α=0.05* para probar si los datos de la tabla están uniformemente distribuidos

α es el nivel de confianza, si *α=0.05* tenemos que el nivel del confianza es de ≈ 95%

$$H_0: \quad R_i \sim \text{Uniform}[0, 1]$$
$$H_1: \quad R_i \nsim \text{Uniform}[0, 1]$$

Ejemplo:
Utiliza la prueba de *chi-cuadrada* con *α=0.05* para probar si los datos de la tabla están uniformemente distribuidos

α es el nivel de confianza, si *α=0.05* tenemos que el nivel del confianza es de ≈ 95%

0.34	0.9	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.7	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.3	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42
0.99	0.17	0.99	0.46	0.05	0.66	0.1	0.42	0.18	0.49
0.37	0.51	0.54	0.01	0.81	0.28	0.69	0.34	0.75	0.49
0.72	0.43	0.56	0.97	0.3	0.94	0.96	0.58	0.73	0.05
0.06	0.39	0.84	0.24	0.4	0.64	0.4	0.19	0.79	0.62
0.18	0.26	0.97	0.88	0.64	0.47	0.6	0.11	0.29	0.78

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Conjunto de datos N = 100
Número de clases o intervalos n= 10

Las clase o intervalos los podemos definirlos como: [0, 0.1), [0.1, 0.2), . . . , [0.9, 1.0).

Con los datos del archivo en Excel calcular el valor de la prueba de X^2

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

0.34	0.9	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.7	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.3	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42
0.99	0.17	0.99	0.46	0.05	0.66	0.1	0.42	0.18	0.49
0.37	0.51	0.54	0.01	0.81	0.28	0.69	0.34	0.75	0.49
0.72	0.43	0.56	0.97	0.3	0.94	0.96	0.58	0.73	0.05
0.06	0.39	0.84	0.24	0.4	0.64	0.4	0.19	0.79	0.62
0.18	0.26	0.97	0.88	0.64	0.47	0.6	0.11	0.29	0.78

Conjunto de datos N = 100

Número de clases o intervalos n= 10

Con los datos del archivo en Excel calcular el valor de la prueba de X^2

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Intervalo	O_i
1	8
2	8
3	10
4	9
5	12
6	8
7	10
8	14
9	10
10	11
	100

Con los datos del archivo en Excel calcular el valor de la prueba de X^2

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Intervalo	O_i	E_i	O_i-E_i	$(O_i-E_i)^2$	$(O_i-E_i)^2/E_i$
1	8				
2	8				
3	10				
4	9				
5	12				
6	8				
7	10				
8	14				
9	10				
10	11				
	100				

Con los datos del archivo en Excel calcular el valor de la prueba de X^2

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Intervalo	O_i	E_i	O_i-E_i	$(O_i-E_i)^2$	$(O_i-E_i)^2/E_i$
1	8	10	-2	4	0.4
2	8	10	-2	4	0.4
3	10	10	0	0	0
4	9	10	-1	1	0.1
5	12	10	2	4	0.4
6	8	10	-2	4	0.4
7	10	10	0	0	0
8	14	10	4	16	1.6
9	10	10	0	0	0
10	11	10	1	1	0.1
	100	100			**3.4**

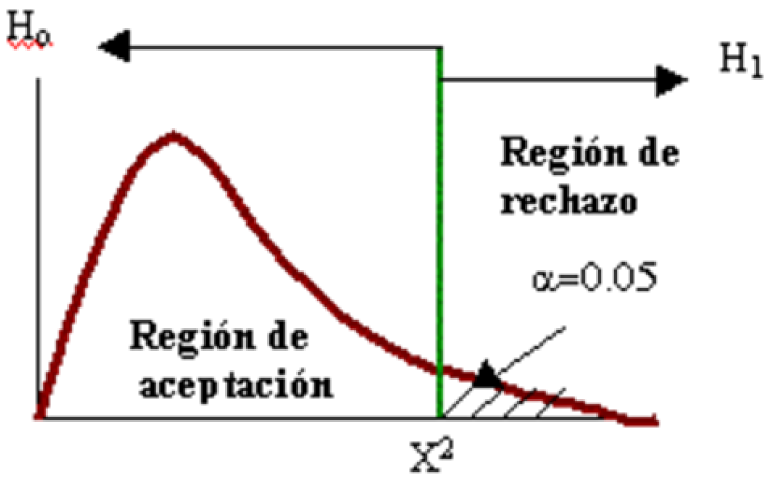
$X^2=3.4$

Con el valor de X^2=3.4 la H_o y con el nivel de confianza *α=0.05 es*

¿aceptada o rechazada?

$$H_0: \quad R_i \sim \text{Uniform}[0, 1]$$
$$H_1: \quad R_i \nsim \text{Uniform}[0, 1]$$

Distribución de *Chi-cuadrada*



Grados de libertad = n-1

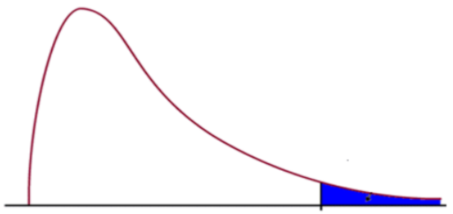
Para el ejemplo con 10 clases tenemos que

Los grados de libertad = 10 – 1 = ***9***

De acuerdo al valor de α=0.05 y grados de libertad = 9, calcula X^2

Distribución Chi-cuadrada

En las columnas se encuentran las áreas bajo la curva a la derecha.



g.l.	$x^2_{0.995}$	$x^2_{0.990}$	$x^2_{0.975}$	$x^2_{0.95}$	$x^2_{0.9}$	$x^2_{0.1}$	$x^2_{0.05}$	$x^2_{0.025}$	$x^2_{0.01}$	$x^2_{0.005}$
1	3.9E-05	0.0002	0.0010	0.0039	0.0158	2.7055	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.6052	5.9915	7.3778	9.2103	10.5966
3	0.0717	0.1148	0.2158	0.3518	0.5844	6.2514	7.8147	9.3484	11.3449	12.8382
4	0.2070	0.2971	0.4844	0.7107	1.0636	7.7794	9.4877	11.1433	13.2767	14.8603
5	0.4117	0.5543	0.8312	1.1455	1.6103	9.2364	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	2.2041	10.6446	12.5916	14.4494	16.8119	18.5476
7	0.9893	1.2390	1.6899	2.1673	2.8331	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	3.4895	13.3616	15.5073	17.5345	20.0902	21.9550
9	1.7349	2.0879	2.7004	3.3251	4.1682	14.6837	16.9190	19.0228	21.6660	23.5894
10	2.1559	2.5582	3.2470	3.9403	4.8652	15.9872	18.3070	20.4832	23.2093	25.1882
11	2.6032	3.0535	3.8157	4.5748	5.5778	17.2750	19.6751	21.9200	24.7250	26.7568

X^2=3.4

La H_o es aceptada, ya que X^2=3.4 es menor que $\chi^2_{0.05,9} = 16.9$

Pruebas de Rachas (aparición aleatoria)

Una ***racha creciente*** es aquella en que un número está seguido de otro mayor
Una ***racha decreciente*** es aquella en que un número está seguido de otro menor

0.00001	0.13154	0.75561	0.45865	0.53277	0.21896	0.04704	0.67886	0.67930	0.93469
0.38350	0.51942	0.83097	0.03457	0.05346	0.52970	0.67115	0.00770	0.38342	0.06684
0.41749	0.68677	0.58898	0.93044	0.84617	0.52693	0.09196	0.65392	0.41600	0.70119
0.91032	0.76220	0.26245	0.04746	0.73608	0.32823	0.63264	0.75641	0.99104	0.36534

N_signos = 39

+	+	-	+	-	-	+	+	+	-
+	+	-	+	+	+	-	+	-	+
+	-	+	-	-	-	+	-	+	+
-	-	-	+	-	+	+	+	-	

Número de rachas
R = 24

R number or runs, R=24

$$\mu_R = E(R) = \frac{2n-1}{3} \qquad \sigma_R^2 = \frac{16n-29}{90}$$

$\mu R = [2(39) - 1]/3 = 25.67$

$\sigma^2 = [16(39) - 29]/90 = 6.61$

$\sigma = 2.57$

$$Z_R = \frac{R-\mu_R}{\sigma_R} \overset{\approx}{\rightsquigarrow} \mathbf{N}(0,1)$$

$Z_R = (R-\mu)/\sigma$

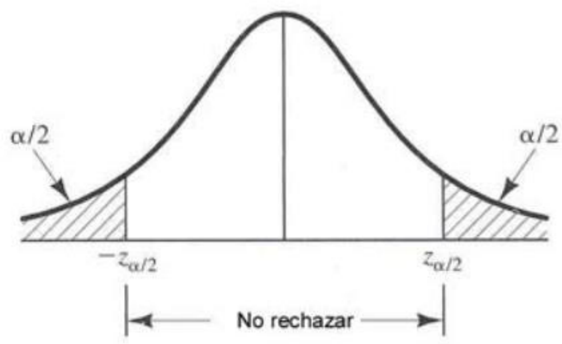
$Z_R = (R-\mu)/\sigma = (24-25.67)/2.57 = -0.65$

Se rechaza la Ho cuando $|Z_R| > z_{\alpha/2}$

Prueba de Rachas Creciente-Decreciente

Se rechaza la H_o si $|Z_R| > z_{\alpha/2}$

Si se requiere 95% de confiabilidad



$1 - \alpha$	$\alpha/2$	$z_{\alpha/2}$
0.90	0.05	1.645
0.95	0.025	1.96
0.99	0.005	2.575

$|Z_R| > z_{\alpha/2}$ ¿$|-0.65| > 1.96$? *FALSO, por lo tanto no se rechaza*