

# Pruebas Estadísticas de Números Pseudoaleatorios

## Prueba de media

Consiste en verificar que los números generados tengan una media estadísticamente igual a  $\frac{1}{2}$ , de esta manera, se analiza la siguiente hipótesis:

$$H_0 : \mu = \frac{1}{2}$$

$$H_1 : \mu \neq \frac{1}{2}$$

Paso 1.

Calcular la media de los  $n$  números generados.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n r_i$$

Paso 2.

Calcular los límites inferior y superior de aceptación:

$$li_{\bar{x}} = \frac{1}{2} - z_{\alpha/2} \left( \frac{1}{12\sqrt{n}} \right) \quad , \quad ls_{\bar{x}} = \frac{1}{2} + z_{\alpha/2} \left( \frac{1}{12\sqrt{n}} \right)$$

Paso 3.

Si el valor de  $\bar{x}$  se encuentra entre  $li_{\bar{x}}$ ,  $ls_{\bar{x}}$ , se acepta que los números tienen una media estadísticamente igual a  $\frac{1}{2}$  con un nivel de confianza del  $1 - \alpha$  %.

**Ejemplo** (Prueba de media). Realice la prueba de media a los siguientes 30 números pseudoaleatorios, con un nivel de confianza del 95%.

0.11352	0.24528	0.12014	0.11963	0.11592	0.66175
0.10047	0.29429	0.41769	0.34661	0.24381	0.66425
0.94441	0.81699	0.82584	0.81231	0.88835	0.71202
0.22591	0.97069	0.33214	0.15361	0.76613	0.42239
0.82506	0.48166	0.56711	0.76950	0.43478	0.22817

Con la ecuación para la media muestral se calcula la media

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n r_i = 0.48735$$

Los límites de aceptación para  $n = 30$  y  $\alpha = 0.05$  son:

$$li_{\bar{x}} = \frac{1}{2} - z_{\alpha/2} \left( \frac{1}{12\sqrt{n}} \right) = \frac{1}{2} - 1.96 \left( \frac{1}{12\sqrt{30}} \right) = 0.4701$$

$$ls_{\bar{x}} = \frac{1}{2} + z_{\alpha/2} \left( \frac{1}{12\sqrt{n}} \right) = \frac{1}{2} + 1.96 \left( \frac{1}{12\sqrt{30}} \right) = 0.5298$$

dado que el valor promedio se encuentra entre los límites, se acepta la hipótesis  $H_0$  es decir, se puede afirmar que la media de los números es estadísticamente igual a  $\frac{1}{2}$ .

# Pruebas Estadísticas de Números Pseudoaleatorios

## Prueba de varianza

Consiste en verificar que los números generados tengan una varianza estadísticamente igual a  $\frac{1}{12} = 0.0833$ , de esta manera, se analiza la siguiente hipótesis:

$$H_0 : \text{Var}(x) = \frac{1}{12}$$

$$H_1 : \text{Var}(x) \neq \frac{1}{12}$$

Paso 1.

Calcular la varianza de los  $n$  números generados.

$$\text{Var}(x) = \frac{\sum_{i=1}^n (r_i - \bar{x})^2}{n-1}$$

Paso 2.

Calcular los límites inferior y superior de aceptación:

$$li_{\text{Var}(x)} = \frac{\chi_{1-\alpha/2, n-1}^2}{12(n-1)}, \quad ls_{\text{Var}(x)} = \frac{\chi_{\alpha/2, n-1}^2}{12(n-1)}$$

Paso 3.

Si el valor de  $\text{Var}(x)$  se encuentra entre  $li_{\text{Var}(x)}$ ,  $ls_{\text{Var}(x)}$ , se acepta que los números tienen una varianza estadísticamente igual a  $\frac{1}{12}$  con un nivel de confianza del  $1 - \alpha$  %.

**Ejemplo** (Prueba de varianza). Realice la prueba de varianza a los siguientes 30 números pseudoaleatorios, con un nivel de confianza del 95%.

0.11352	0.24528	0.12014	0.11963	0.11592	0.66175
0.10047	0.29429	0.41769	0.34661	0.24381	0.66425
0.94441	0.81699	0.82584	0.81231	0.88835	0.71202
0.22591	0.97069	0.33214	0.15361	0.76613	0.42239
0.82506	0.48166	0.56711	0.76950	0.43478	0.22817

Con la ecuación para la varianza muestral se calcula:

$$Var(x) = \frac{\sum_{i=1}^n (r_i - \bar{x})^2}{n-1} = 0.0854$$

Los límites de aceptación para  $n = 30$  y  $\alpha = 0.05$  son:

$$li_{Var(x)} = \frac{\chi_{1-\alpha/2, n-1}^2}{12(n-1)} = \frac{\chi_{0.975, 29}^2}{12(29)} = \frac{16.04}{348} = 0.0461$$

$$ls_{Var(x)} = \frac{\chi_{\alpha/2, n-1}^2}{12(n-1)} = \frac{\chi_{0.025, 29}^2}{12(29)} = \frac{45.7}{348} = 0.1313$$

dado que el valor de la varianza muestral se encuentra entre los límites, se acepta la hipótesis  $H_0$  es decir, se puede afirmar que la varianza de los números es estadísticamente igual a  $\frac{1}{12}$ .

# Pruebas Estadísticas de Números Pseudoaleatorios

## Prueba de forma

Para realizar esta prueba se utiliza la prueba de bondad de ajuste  $\chi^2$ , ya descrita anteriormente. Esta prueba se empleará específicamente para números aleatorios uniformes entre 0 y 1, para probar que un conjunto de datos sigue esta distribución. De esta manera la hipótesis propuesta se resume como sigue:

$$H_0 : r_i \sim U[0,1]$$

$$H_1 : r_i \sim \neq U[0,1]$$

**Ejemplo.** Tomando los 30 números del ejemplo anterior, determine con un nivel de confianza del 95% si pertenecen a una población uniforme. Dividiendo el rango de 0 a 1 en 5 intervalos y clasificando los 30 números según su valor se obtiene la siguiente tabla:

Intervalo	FO	FE=30/5
0.0 - 0.1	6	6
0.2 - 0.4	7	6
0.4 - 0.6	5	6
0.6 - 0.8	5	6
0.8 - 1.0	7	6

Se calcula  $C$  a partir de la ecuación

$$C = \sum_{i=1}^m \frac{(FE_i - FO_i)^2}{FE_i}$$

se obtiene un valor de  $C = 0.6667$ . Se compara con el valor de tablas  $\chi^2$  con  $5-1=4$  grados de libertad y un nivel de 5% que es igual a 9.49 y la comparación indica que los números generados siguen una distribución uniforme entre 0 y 1.

# Pruebas Estadísticas de Números Pseudoaleatorios

## Prueba de independencia

Las pruebas de independencia consisten en demostrar que los números generados son estadísticamente independientes entre sí, esto es, que no dependen uno del otro. Para esto se propone la siguiente hipótesis:

$$H_0 : r_i \sim \text{Independiente}$$

$$H_1 : r_i \sim \text{Dependiente}$$

Para realizar estas pruebas de hipótesis existen varios métodos, puede seleccionarse cualquiera de la siguiente lista:

- Prueba de póker.
- Prueba de corridas arriba y abajo.
- Prueba de corridas arriba y debajo de la media.
- Prueba de la longitud de las corridas.
- Prueba de distancia.
- Prueba de series.
- Prueba de huecos.

# Prueba de independencia

## Prueba de Póker

$H_0 : r_i \sim$  Independiente

$H_1 : r_i \sim$  Dependiente

**Paso 1.** Calcular las probabilidades esperadas para un juego de póker con 5 cartas numeradas del 0 al 9 con reemplazo, se tienen 7 eventos o intervalos, con las siguientes probabilidades:

$$P(\text{Pachuca}) = 0.3024 \quad P(\text{Un par}) = 0.5040$$

$$P(\text{Dos pares}) = 0.1080 \quad P(\text{Una tercia}) = 0.0720$$

$$P(\text{Full}) = 0.0090 \quad P(\text{Póker}) = 0.0045$$

$$P(\text{Quintilla}) = 0.0001$$

**Paso 2.** Calcular la frecuencia esperada de cada uno de los eventos ( $FE_i$ ) multiplicando la probabilidad de cada evento por el número de números aleatorios generados.

**Paso 3.** Para cada número aleatorio generado verificar (imaginando que es una mano de póker) si es pachuca, un par, dos pares, etc., tomando los primeros cinco dígitos a la derecha del punto decimal. Por ejemplo, 0.3408 es un par, 0.44343 es un full, 0.00323 dos pares, etc. Con estos resultados se genera una tabla de frecuencias de estos eventos. La frecuencia observada de cada uno de los eventos se denota como ( $FO_i$ ).

**Paso 4.** Calcular el estadístico  $C$  con  $m = 7$ .

**Paso 5.** Si el valor de  $C$  es menor o igual al estadístico de tablas  $\chi^2$  con 6 grados de libertad y una probabilidad de rechazo  $\alpha$ , entonces se acepta que los datos son estadísticamente independientes entre sí.

**Ejemplo.** Realice la prueba de póker a los siguientes 30 números con un nivel de confianza del 95%.

0.72484 0.48999 0.50502 0.39528 0.36782 0.90234  
 0.71890 0.61234 0.86322 0.94134 0.99872 0.27657  
 0.34565 0.02345 0.67347 0.10987 0.25678 0.25593  
 0.82345 0.12387 0.05389 0.82474 0.59289 0.36782  
 0.03991 0.10461 0.93716 0.16894 0.98953 0.73231

Agrupando los números de acuerdo con sus dígitos, como si fuera una mano de póker se obtiene la siguiente tabla de frecuencias:

Intervalo	FO	PE	FE=( $n*PE$ )
Pachuca	14	0.3024	9.072
Un par	15	0.5040	15.120
Dos pares	1	0.1080	3.240
Una tercia	1	0.0720	2.160
Full	0	0.0090	0.270
Póker	0	0.0045	0.135
Quintilla	0	0.0001	0.003

El cálculo de  $C$  es igual a 4.25 que comparado contra el valor de tablas  $\chi^2$  con  $7-1 = 6$  grados de libertad, y con un nivel de 5% que es igual a 12.59, indica que los números generados son estadísticamente independientes.

# Prueba de independencia

## Prueba de Corridas

$$H_0 : r_i \sim \text{Independiente}$$

$$H_1 : r_i \sim \text{Dependiente}$$

**Paso 1.** Clasificar cada número aleatorio con respecto al anterior, de acuerdo con:

$$\text{Si } r_i \leq r_{i-1} \quad r_i = -$$

$$\text{Si } r_i > r_{i-1} \quad r_i = +$$

**Paso 2.** Calcular el número de corridas observadas  $h$ . Una corrida se forma por un conjunto de números aleatorios consecutivos del mismo signo.

**Paso 3.** Calcular  $E(h)$  y  $\text{Var}(h)$  de acuerdo con:

$$E(h) = \frac{2n-1}{3}$$

$$\text{Var}(h) = \frac{16n-29}{90}$$

donde  $n$  es el número de datos generados.

**Paso 4.** Calcular el estadístico  $z = \frac{(h - E(h))}{\sqrt{\text{Var}(h)}}$ , si el valor absoluto

$|z|$  es menor que el valor crítico  $Z_{\alpha/2}$  se acepta la hipótesis de independencia.

**Ejemplo.** Determine si la siguiente serie de 30 números puede ser aceptada como independiente con un nivel de confianza del 95%, usando la prueba de corridas

0.72484	0.48999	0.50502	0.39528	0.36782	0.90234
0.71890	0.61234	0.86322	0.94134	0.99872	0.27657
0.34565	0.02345	0.67347	0.10987	0.25678	0.25593
0.82345	0.12387	0.05389	0.82474	0.59289	0.36782
0.03991	0.10461	0.93716	0.16894	0.98953	0.73231

La secuencia de corridas es:

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

donde existen  $h = 21$  corridas. Con  $n = 30$  tenemos de las ecuaciones anteriores que  $E(h) = 19.667$  corridas y  $Var(h) = 5.01$ .

El valor de  $z = \frac{(h - E(h))}{\sqrt{Var(h)}} = \frac{21 - 19.667}{\sqrt{5.01}} = 0.5955$

Comparado contra  $Z_{0.025} = 1.96$ , de tal manera que indica que los números generados son estadísticamente independientes.

# Prueba de independencia

## Prueba de Series

$H_0 : r_i \sim$  Independiente

$H_1 : r_i \sim$  Dependiente

### Paso 1.

Crear un histograma de dos dimensiones con  $m$  intervalos, clasificando cada pareja de números consecutivos  $(r_i, r_{i+1})$  dentro de las casillas de dicho histograma de frecuencias. El número total de pares ordenados en cada casilla formará la frecuencia observada:  $FO_i$ .

### Paso 2.

Calcular la frecuencia esperada en cada casilla  $FE_i$  de acuerdo con  $FE_i = \text{núm}/m$  donde  $\text{núm}$  es el número total de parejas ordenadas.

### Paso 3.

Calcular el valor del estadístico  $C$ .

### Paso 4.

Si el valor de  $C$  es menor o igual al valor en tablas  $\chi^2$  con  $m-1$  grados de libertad y una probabilidad de rechazo  $\alpha$ , entonces se acepta la hipótesis de independencia.

**Ejemplo.** Determine si la siguiente serie de 30 números puede ser aceptada como independiente con un nivel de confianza del 95%, usando la prueba de series

|         |         |         |         |         |         |
|---------|---------|---------|---------|---------|---------|
| 0.72484 | 0.48999 | 0.50502 | 0.39528 | 0.36782 | 0.90234 |
| 0.71890 | 0.61234 | 0.86322 | 0.94134 | 0.99872 | 0.27657 |
| 0.34565 | 0.02345 | 0.67347 | 0.10987 | 0.25678 | 0.25593 |
| 0.82345 | 0.12387 | 0.05389 | 0.82474 | 0.59289 | 0.36782 |
| 0.03991 | 0.10461 | 0.93716 | 0.16894 | 0.98953 | 0.73231 |

Al formar parejas ordenadas se obtiene:

(.72484, .48999) (.48999, .50502) (.50502, .39528) ... (.98953, .73231).

La clasificación en una tabla de frecuencias de dos dimensiones de  $4 \times 4$ , ( $m=16$ ), queda:

|           |      |      |      |      |   |       |
|-----------|------|------|------|------|---|-------|
| $r_{i+1}$ | 1    | 3    | 2    | 1    | 2 |       |
|           | 0.75 | 1    | 1    | 1    | 3 |       |
|           | 0.50 | 1    | 3    | 3    | 1 |       |
|           | 0.25 | 2    | 2    | 1    | 2 |       |
|           | 0    | 0.25 | 0.50 | 0.75 | 1 | $r_i$ |

Tomando en cuenta que se tienen 29 parejas ordenadas clasificadas uniformemente en 16 casillas, la frecuencia esperada FE en cada una es 1.8125 y al calcular el estadístico C, para cada una de las 16 celdas o intervalos de la tabla anterior, se tiene:

$$C = \sum_{i=1}^m \frac{(FE_i - FO_i)^2}{FE_i} = \sum_{i=1}^{16} \frac{(1.8125 - FO_i)^2}{1.8125}$$

$$C = \frac{1}{1.8125} [7(1.8125 - 1)^2 + 5(1.8125 - 2)^2 + 4(1.8125 - 3)^2] = 5.75$$

El valor de la tabla  $\chi^2$  con un nivel de confianza del 95% y con 15 grados de libertad es igual a 25. Si se compara  $C = 5.75$  con este valor, se acepta la independencia de la serie de números.