

COMPARACION POR SIMULACION DE TRES METODOS
DE ESTIMACION DE PARAMETROS PARA LA DISTRI
BUCION DE GUMBELL CON MUESTRAS ALEATORIAS
Y AUTOCORRELACIONADAS

Pedro Nel Pacheco Durán

Profesor Asociado
Universidad Nacional

Ricardo Cogua Suárez

Estadístico
Universidad Nacional

Resumen: En este trabajo se muestra una comparación del comportamiento de los estimadores de momentos, máximo verosímiles y de momentos de probabilidad ponderados para los parámetros y percentiles de la distribución de Gumbell, por medio de la simulación estadística bajo dos supuestos.

1. Muestras aleatorias.
2. Muestras autocorrelacionadas de primer orden que se asumen aleatorias.

Para realizar las simulaciones a gran escala se implementó un programa en lenguas SAS en el computador del centro de cálculo de la Universidad Nacional. De los resultados

obtenidos con dicho programa se concluye que el mejor método para la estimación de parámetros y percentiles es el de momentos de probabilidad ponderados, ya que para todo tamaño de muestra es prácticamente insesgado y la varianza es comparable con la de los otros métodos para la mayoría de los tamaños de muestra.

1. Diseño experimental.

1.1. Generalidades.

El proceso en términos generales consiste en generar muestras aleatorias de diferentes tamaños de la distribución de Gumbell y, para los mismos tamaños de muestra, generar valores de un proceso autocorrelacionado de orden uno con distribución marginal Gumbell.

La implementación de este procedimiento requiere entonces de un mecanismo para generar valores que se distribuyan Gumbell, y es aquí donde se puede aprovechar que la función de distribución es inversible. Más formalmente: Una variable aleatoria X se dice que tiene distribución de Gumbell si

$$F(x) = \text{Exp}\{-\text{Exp}[-((x-\mu)/\alpha)]\}$$

donde $F(x) = P[X \leq x]$.

La inversa de la distribución de Gumbell es:

$$x = \mu - \alpha \ln [-\ln(F)]$$

y de la teoría estadística se conoce que $F(x)$ tiene distribución uniforme en el intervalo unitario. Por lo tanto, basta

generar un valor $U(0,1)$ y aplicar la función inversa de F para generar un valor Gumbell.

Se consideró adecuado generar muestras de tamaño $n = 5, 10, 20, 30, 50, 80, 100$ y 1000 , ya que este rango cubre la gran mayoría de los tamaños de muestra posibles en estudios aplicados y también los tres tipos de tamaño muestral (grande, mediano y pequeño), de acuerdo con los teoremas límites.

Una vez generadas las muestras, se calculan las estimaciones de los parámetros de la distribución (α y μ), y a partir de estos los percentiles $x(F)$ donde $F \in \Omega$:

$$\Omega = \{0.001, 0.010, 0.020, 0.050, 0.100, 0.250, 0.500, 0.750, 0.900, 0.950, 0.980, 0.990, 0.999\}.$$

Con este procedimiento se determinan para cada muestra las estimaciones de los parámetros y percentiles utilizando

- 1 Método de momentos convencionales.
- 2 Método de máxima verosimilitud.
- 3 Método de momentos de probabilidad ponderados.

Ahora, si $\hat{\theta}$ denota un estimador de $\theta \in [\alpha \ \mu \ x(F)]$, tomando como base, N muestras de tamaño n : la media y la varianza de los $\hat{\theta}$ se obtiene como:

$$\hat{\mu}[\hat{\theta}] = \left(\sum_{i=1}^N \hat{\theta}_i \right) / N$$

$$\hat{\sigma}[\hat{\theta}] = \left\{ 1/N \sum_{i=1}^N \hat{\theta}_i^2 - \left(1/N \sum_{i=1}^N \hat{\theta}_i \right)^2 \right\}$$

respectivamente, donde $N = 50000$ si $n \leq 100$ y $N = 2000$, si

$n > 100$.

Para cada uno de los tres métodos y para cada muestra de tamaño n , el sesgo de los estimadores de θ se calcula así:

$$\tilde{\mu} [\theta - \hat{\theta}] = \theta - \tilde{\mu}[\hat{\theta}]$$

y el error cuadrático medio como:

$$\text{ECM}(\hat{\theta}/n, \text{método}) = \tilde{\mu}^2[\theta - \hat{\theta}] + \tilde{\sigma}^2[\hat{\theta}]$$

La eficacia relativa de dos métodos en la estimación de θ da do un tamaño de muestra n se define como:

$$\text{EF}[\hat{\theta}, \text{Método 1}, \text{Método 2}, n] = \frac{\text{ECM}(\hat{\theta}, \text{Método 1}, n)}{\text{ECM}(\hat{\theta}, \text{Método 2}, n)}$$

Si $\text{EF} < 1$, el Método 1 es más eficiente en la estimación de $\hat{\theta}$ que el Método 2. Si $\text{EF} > 1$ el Método 2 es más eficiente en la estimación de θ que el Método 1. Si $\text{EF} = 1$ no se puede es tablecer una preferencia con base en la eficiencia del Método 1 relativo al Método 2.

Para cada uno de los tres métodos se consideraron dos casos:

1. Una muestra de tamaño n que proviene de un proceso puramente aleatorio.
2. Una muestra de tamaño n que procede de un proceso serialmente correlacionado visto como puramente aleatorio.

1.2. Generación de números aleatorios Gumbell.

Generación de muestras aleatorias.

Para el caso de un proceso aleatorio las N muestras de tamaño n fueron generadas usando la forma inversa de la distribución de Gumbell. Se realizó un programa diseñado en el paquete SAS tomando un valor numérico como semilla, usando la función RANUNI que genera números aleatorios $U(0,1)$ de un generador multiplicativo para un módulo primo igual a $2^{31}-1$. La semilla debe ser una constante numérica que es menor que $2^{31}-1$ que se tomó como la fecha de ejecución del programa.

Generación de un Proceso autocorrelacionado.

Para el proceso serialmente correlacionado los datos fueron generados usando el método sugerido por H.A. THOMAS, (1976) así:

$$Z_i = P_Z Z_{i-1} + (1 + P_Z^2)^{\frac{1}{2}} \delta_1$$

que representa un proceso de Markov, donde la correlación de primer orden de los Z_1 y δ_1 es una variable aleatoria normalmente distribuida con media cero y varianza uno, independiente de Z_{t-1} (δ_1 son generados mediante la función RANNOR). De la misma manera Z_1 esta normalmente distribuida con media cero y varianza uno.

Dando los valores de Z_1 , los correspondientes valores de $F(Z_1) = P[Z \leq z_1]$ estan dados por:

$$F(z_i) = \int_{-\infty}^{z_1} \delta_N(z) dz ,$$

donde δ_N es la función de densidad para la distribución normal. Los valores de X_i serialmente correlacionados para la distribución de Gumbell fueron obtenidos así:

$$X_i = \mu - \alpha \ln \{- \ln (F(z_i)) \}$$

Los valores de $\delta_1 = 1, 2, \dots, n$ fueron generados aplicando la técnica de BOX y MULLER (1958) de la función RANOR de SAS. El valor de P_z se tomó como 0.5, lo suficientemente grande para asegurar el efecto de la correlación serial.

2. Marco teórico.

Momentos de probabilidad ponderados.

Greenwood y otros (1979) definieron momentos de probabilidad ponderados así:

$$M_{i,j,k} = E[X^i F^j (1-F)^k]$$

donde $i, j, k \in R$. Si i, j, k son enteros no negativos, $M_{i,j,k}$ es proporcional al momento i -ésimo respecto al origen de la estadística de orden $(j+1)$ -ésima, de una muestra de tamaño $(k+j+1)$. En especial:

$$M_{i,j,k} = B[j+1, k+1] E[X_{(j+1), (k+j+1)}^i]$$

donde $B[,]$ corresponde a la función Beta.

Para $j = 0$ y $1 = k$, se escribirá:

$$M_{(k)} = M_{1,0,k} = (1/k+1) E [X_{1,k+1}] \quad (1)$$

se probó que $M_{(k)}$ es un estimador insesgado de $M_{(k)}$ para una muestra aleatoria de tamaño n , donde k es un entero no negativo. $M_{(k)}$ se obtiene como sigue: Supongamos que las n observaciones están ordenadas en forma ascendente desde X_1 hasta X_n .

El número total de posibles formas de obtener un conjunto de submuestras de tamaño $(k+1)$ es:

$$N = \binom{n}{k+1}$$

el número de estas N submuestras que contienen x_1 , la observación más pequeña es:

$$N_1 = \binom{n-1}{k}$$

similarmente, para las N submuestras que contienen x_2 pero no x_1 , es decir, el número de submuestras posibles en los cuales x_2 es la observación más pequeña es:

$$N_2 = \binom{n-2}{k}$$

En general el número de submuestras de tamaño $(k+1)$ que contienen X_i pero ninguna de las anteriores (X_1, \dots, X_{i-1}) , es

$$N_i = \binom{n-i}{k}$$

Nótese que si $i \geq n - k + 1$, $N_i = 0$.

Dado que existe un total de N posibles submuestras, de N_1 , en los cuales X_i puede ser la observación más pequeña, en un estimador de probabilidad de que x_i sea la observación más pequeña en una submuestra de tamaño $(k+1)$ tomada de una muestra de tamaño n es:

$$\left(u \left| \begin{array}{c} k \\ i \end{array} \right. n \right) = \frac{N_i}{N} = \binom{N-i}{k} / \binom{n}{k+1}$$

Note que

$$\left(u \left| \begin{array}{c} k \\ i \end{array} \right. n \right)$$

es un estimador no paramétrico y cuya distribución no depende de la densidad de la variable involucrada. Por lo tanto, un estimador del mínimo de una submuestra de tamaño $(k+1)$ extraída de una muestra de tamaño n es:

$$E[X_{1,k+1} / n] = \sum_{i=1}^n x_i \binom{n-i}{k} / \binom{n}{k+1}$$

pero de la relación establecida en (1) se obtiene

$$\begin{aligned} M_{(k)} &= 1 / (k+1) \sum_{i=1}^n x_i \binom{n-i}{k} / \binom{n}{k+1} \\ &= 1 / n \sum_{i=1}^n x_i \binom{n-i}{k} / \binom{n+1}{k+1} \\ &= 1 / n \sum_{i=1}^{n-k} x_i \binom{n-i}{k} / \binom{n-1}{k} \end{aligned}$$

El valor esperado de X_i viene dado por:

$$E[X] = i \binom{n}{i} \int_0^1 0 F^{i-1} (1-F)^{n-1} dF$$

donde $0 \equiv 0(F)$ y $F \equiv F(0)$, sí el valor esperado $M_{(k)}$ puede ser expresado como:

$$E[M_{(k)}] = 1/n \sum_{i=1}^{n-k} \left[\binom{n-i}{k} / \binom{n-1}{k} \right] E[X_i]$$

$$E[M_{(k)}] = \int_0^1 0 \left[\sum_{j=0}^{n-1-k} \binom{n-1-k}{j} F^j (1-F)^{n-j-k-1} \right] \cdot (1-F)^k dF .$$

Donde $j = i - 1$. Esto denota que:

$$\begin{aligned} & \sum_{j=0}^{n-1-k} \binom{n-1-k}{j} F^j (1-F)^{n-j-k-1} \\ &= [F + (1-F)]^{n-k-1} = 1 , \end{aligned}$$

donde

$$E[M_{(k)}] = \int_0^1 0 (1-F)^k dF = M_{(k)} .$$

De donde se concluye que este estimador de $M_{(k)}$ es insesgado.

A partir de este resultado, se pueden obtener estimadores de los parámetros de la distribución de Gumbell a partir de $M_{(k)}$ como sigue:

$$\alpha = [M_{(k)} - 2M_{(1)}] / 1n(2)$$

$$\mu = M_{(0)} - \varepsilon \alpha_k .$$

3. Análisis de resultados.

3.1. Parámetro de localización para muestras aleatorias.

3.1.1. Sesgo.

Como podemos observar en forma clara en la figura (2), Tabla (1) se presentó una jerarquía. El método de momentos es el más sesgado, seguido por el método de máxima verosimilitud. El método de momentos de probabilidad ponderados presenta un insesgamiento muy notorio para todos los tamaños de muestra examinados, por supuesto es el mejor de los tres.

En los tres métodos podemos ver claramente la tendencia a sobreestimar el parámetro.

Los tres métodos pueden ser comparados en tamaños de muestra grandes.

3.1.2. Varianza.

Como podemos observar en la figura (4) Tabla (1), la varianza disminuye cuando n aumenta.

El método de momentos de probabilidad ponderados presenta varianza mínima para todo n , seguida del método de máxi

ma verosimilitud y momentos convencional.

Para todos los tamaños de muestras se observa que momentos de probabilidad ponderados presenta varianza mínima, aunque cuando n es mayor de 20 las varianzas son practicamente iguales.

3.1.3. Error cuadrático medio.

En la figura (6) Tabla (1), se puede observar que el mejor método es momentos de probabilidad ponderados seguido por máxima verosimilitud y momentos convencionales.

3.2. Parámetro de dispersión para muestras aleatorias.

3.2.1. Sesgo.

Como se observa en la figura (1) Tabla (1), los tres métodos presentan una subestimación $\alpha \leq 1$.

Al igual que en los casos anteriores en la figura (1) Tabla (1) se observa que el sesgo decrece exponencialmente para todos los tamaños de muestra analizados en los tres mé todos.

Podemos ver a lo largo de todos los tamaños de muestra analizados una jerarquía muy notoria, momentos de probabilidad ponderados es el más insesgado, seguido por máxima verosimilitud y momentos convencionales.

El método de momentos de probabilidad ponderado es muy bueno para muestras pequeñas y practicamente insesgado para tamaño de muestra superior a 10.

3.2.2. Varianza.

El estimador máximo verosímil es el que presenta varianza mínima para todos los tamaños de muestra como lo podemos ver en la figura (3) Tabla (1).

Para los métodos de momento y momentos de probabilidad ponderados la varianza está en función de n .

Para un n pequeño ($n = 5, 10$), momentos convencionales es mayor o igual que momentos de probabilidad.

Para un $n > 20$ momentos de probabilidad ponderados es mayor que momentos convencionales.

Se observa que a medida que aumenta el tamaño de muestra, la diferencia que se presenta entre las varianzas del método de máxima verosimilitud y el método de momentos de probabilidad ponderados es cada vez más pequeño, hasta ser aproximadamente nula a partir de un n mayor que 50. La varianza de momentos es muy grande en comparación con las anteriores.

3.2.3. Error cuadrático medio.

En la figura (5) Tabla (1), podemos ver que el método de máxima verosimilitud es el mejor para todo n .

Para los métodos de momentos y momentos de probabilidad ponderados el error cuadrático medio esta en función de n .

Para muestras pequeñas ($n = 5$), momentos convencionales es mayor o igual que momentos de probabilidad ponderados.

Para muestras en las cuales ($n > 10$), momentos de probabilidad ponderados es mayor que momentos convencionales.

El error cuadrático medio de máxima verosimilitud es aproximadamente 0.75 del de momentos de probabilidad ponderados.

Para alfa el cambio entre error sistemático (sesgo) y error aleatorio (varianza) resulta desfavorable en términos del error cuadrático medio. Sin embargo es mejor en general el método de momentos de probabilidad ponderado por insesgamiento mientras que el método de máxima verosimilitud por varianza.

Por la utilización de computadores es más fácil trabajar con momentos de probabilidad ponderados, pero con la pro fusión de computadores y programas, los dos métodos son igual mente utilizables.

3.3. Percentiles para muestras aleatorias.

3.3.1. Sesgo.

Como se puede observar en la Tabla (2) existe una jerarquía muy clara. El método de momentos de probabilidad pon derados es el mejor de los tres métodos para todo n , el método de máxima verosimilitud se encuentra en el medio y el peor de todos es el método de momentos convencionales.

Para los métodos de momentos de probabilidad ponderados y máxima verosimilitud el sesgo es muy similar en muestras de tamaño mayores o iguales a 100.

Dentro del rango medio (0.25 - 0.75) los percentiles se estiman mucho mejor que para los demás.

3.3.2. Varianza.

Como se observa en la Tabla (5) en los percentiles del rango medio, la varianza de momentos de probabilidad ponderados es menor que la de máxima verosimilitud y la de momentos convencionales.

En los extremos de la Tabla (4) se observa que la varianza de máxima verosimilitud es la menor, pero para muestras grandes, los métodos de momentos de probabilidad ponderados y máxima verosimilitud son comprobables.

Para todos los tamaños de muestra analizados, el método de momentos posee la varianza más grande.

3.3.3. Error cuadrático medio.

En la Tabla (5) podemos darnos cuenta que para el rango medio (0.25 - 0.75) es el método de momentos de probabilidad ponderados es el mejor, seguido por máxima verosimilitud.

4. Conclusiones.

La comparación de un nuevo método de estimación, el método de momentos de probabilidad ponderados con otros dos métodos más antiguos, máxima verosimilitud y momentos convencionales fue explorado mediante un programa en el paquete

SAS. Los experimentos aportaron la estimación de los parámetros y cuantiles de un proceso puramente aleatorio y un proceso autocorrelacionado tratado como puramente aleatorio.

Las conclusiones derivadas de estos experimentos son:

4.1. Para un proceso puramente aleatorio.

El sesgo de momentos de probabilidad ponderados es siempre el más pequeño para todos los tamaños de muestra estudiados. Se nota que desde muestras de tamaño pequeño es casi despreciable y a medida que n aumenta. Las estimaciones de máxima verosimilitud se acercan a las de momentos de probabilidad ponderada.

El método de momentos de probabilidad ponderados cambia una gran cantidad de error sistemático por error aleatorio para tamaños de muestra muy pequeños ($n = 5, 10, 20$) y, en general; para los percentiles que están fuera de la banda central (0,25 - 0,75). En otros casos este cambio no afecta prácticamente en nada y resulta mucho mejor momentos de probabilidad ponderado que máxima verosimilitud. Es decir: si el sesgo es la principal preocupación el mejor en todos los casos es el método de momentos de probabilidad ponderados. En el supuesto que el sesgo sea secundario será mejor usar el método de máxima verosimilitud, cuando se estima con tamaños de muestra pequeña ($n = 5, 10, 20$) o para percentiles que se encuentran en la franja externa (0,25 - 0,75).

En ningún caso se recomienda el uso del método de momentos convencionales.

4.2. Para un proceso autocorrelacionado.

Para el proceso autocorrelacionado visto como puramente aleatorio se conservan los mismos resultados obtenidos en el caso anterior, con la diferencia de que presenta mayor sesgo y mayor varianza. Sin embargo, se puede notar que los estimadores son asintóticamente insesgados pero con velocidad de convergencia menor a la de una muestra aleatoria. Por último, si se sabe que las observaciones provienen de un proceso autocorrelacionado, basta con tener una muestra de tamaño moderado para garantizar buenos resultados. En el caso de muestras pequeñas se recomienda tener mucho cuidado, porque los errores (sesgo y varianza) pueden ser demasiado grandes dependiendo del área de aplicación.

*

ESTIMACION DE PARAMETROS Y PROPIEDADES ESTADISTICAS PARA MUESTRAS ALEATORIAS

METODO DE MUESTRA	TAMANO DE MUESTRA	ALEA			MIU				
		SESGO	VARIANZA	ECM	EFICIENCIA	SESGO	VARIANZA	ECM	EFICIENCIA
MO	5	-0.188	0.135	0.170	0.892	-0.110	0.244	0.256	1.090
MV		-0.159	0.117	0.142	1.195	-0.085	0.240	0.247	1.036
MP		-0.003	0.191	0.191	0.746	-0.003	0.235	0.235	1.052
MO	10	-0.094	0.083	0.092	1.038	-0.057	0.117	0.120	1.066
MV		-0.076	0.061	0.067	1.370	-0.041	0.114	0.115	1.041
MP		-0.002	0.089	0.089	0.758	-0.002	0.113	0.113	1.024
MO	20	-0.050	0.046	0.049	1.171	-0.028	0.058	0.059	1.053
MV		-0.039	0.030	0.032	1.530	-0.019	0.056	0.056	1.044
MP		0.000	0.042	0.042	0.766	0.000	0.056	0.056	1.009
MO	30	-0.033	0.032	0.033	1.216	-0.019	0.039	0.039	1.046
MV		-0.025	0.020	0.021	1.598	-0.012	0.037	0.038	1.041
MP		0.000	0.028	0.028	0.761	0.001	0.037	0.037	1.005
MO	50	-0.019	0.020	0.021	1.265	-0.011	0.023	0.023	1.049
MV		-0.015	0.012	0.012	1.662	-0.007	0.022	0.022	1.048
MP		-0.001	0.016	0.016	0.761	0.000	0.022	0.022	1.001
MO	80	-0.014	0.013	0.013	1.304	-0.008	0.014	0.015	1.046
MV		-0.010	0.008	0.008	1.715	-0.005	0.014	0.014	1.046
MP		-0.001	0.010	0.010	0.760	0.000	0.014	0.014	0.999
MO	100	-0.010	0.011	0.011	1.315	-0.006	0.012	0.012	1.048
MV		-0.007	0.006	0.006	1.737	-0.003	0.011	0.011	1.050
MP		0.000	0.008	0.008	0.757	0.000	0.011	0.011	0.998
MO	1000	-0.001	0.001	0.001	1.362	0.000	0.001	0.001	1.049
MV		-0.001	0.001	0.001	1.757	0.000	0.001	0.001	1.052
MP		0.000	0.001	0.001	0.775	0.001	0.001	0.001	0.997

TABLA 1

TABLA 2

ESTIMACION DE PERCENTILES PARA MUESTRAS CON CORRELACION CERO

METODO	TAMANO DE MUESTRA	ESTIMACION DE PERCENTILES PARA MUESTRAS CON CORRELACION CERO													
		F=0.001	F=0.010	F=0.020	F=0.050	F=0.100	F=0.250	F=0.500	F=0.750	F=0.900	F=0.950	F=0.980	F=0.990		
NO	5	NO	-1.461	-1.132	-0.989	-0.782	-0.568	-0.156	0.408	1.123	1.939	2.524	3.381	3.849	5.724
		NP	-1.540	-1.200	-1.062	-0.838	-0.617	-0.190	0.393	1.132	1.977	2.562	3.365	3.952	5.892
		HP	-1.825	-1.520	-1.357	-1.091	-0.829	-0.323	0.369	1.246	2.248	2.966	3.895	4.592	6.893
NO	10	NO	-1.693	-1.326	-1.178	-0.937	-0.688	-0.238	0.390	1.186	2.096	2.748	3.592	4.225	6.315
		NP	-1.745	-1.371	-1.220	-0.973	-0.730	-0.261	0.380	1.193	2.121	2.786	3.648	4.293	6.425
		HP	-1.935	-1.528	-1.365	-1.098	-0.834	-0.325	0.369	1.250	2.257	2.978	3.912	4.612	6.924
NO	20	NO	-1.808	-1.423	-1.268	-1.014	-0.764	-0.282	0.376	1.212	2.166	2.850	3.735	4.399	6.591
		NP	-1.839	-1.449	-1.292	-1.036	-0.783	-0.295	0.371	1.216	2.181	2.873	3.768	4.439	6.656
		HP	-1.932	-1.527	-1.364	-1.097	-0.834	-0.327	0.366	1.245	2.249	2.969	3.900	4.598	6.904
NO	30	NO	-1.850	-1.458	-1.300	-1.042	-0.788	-0.297	0.373	1.223	2.194	2.890	3.790	4.465	6.695
		NP	-1.872	-1.477	-1.318	-1.058	-0.801	-0.307	0.369	1.226	2.205	2.907	3.815	4.496	6.744
		HP	-1.934	-1.529	-1.365	-1.098	-0.835	-0.328	0.366	1.245	2.250	2.970	3.902	4.601	6.909
NO	50	NO	-1.884	-1.486	-1.326	-1.065	-0.807	-0.309	0.371	1.233	2.218	2.924	3.838	4.522	6.785
		NP	-1.896	-1.497	-1.336	-1.073	-0.814	-0.315	0.368	1.234	2.223	2.932	3.850	4.538	6.810
		HP	-1.935	-1.529	-1.366	-1.099	-0.835	-0.327	0.366	1.247	2.252	2.972	3.905	4.604	6.913
NO	80	NO	-1.898	-1.498	-1.337	-1.074	-0.815	-0.314	0.369	1.237	2.227	2.937	3.856	4.544	6.820
		NP	-1.908	-1.507	-1.345	-1.081	-0.820	-0.319	0.368	1.238	2.232	2.945	3.867	4.558	6.841
		HP	-1.931	-1.525	-1.362	-1.096	-0.833	-0.326	0.367	1.245	2.249	2.968	3.899	4.597	6.902
NO	100	NO	-1.907	-1.506	-1.345	-1.080	-0.820	-0.318	0.368	1.239	2.233	2.945	3.868	4.559	6.842
		NP	-1.915	-1.513	-1.351	-1.086	-0.825	-0.321	0.367	1.240	2.237	2.951	3.876	4.569	6.859
		HP	-1.934	-1.528	-1.365	-1.098	-0.835	-0.327	0.366	1.246	2.250	2.970	3.902	4.601	6.908
NO	1000	NO	-1.930	-1.525	-1.362	-1.096	-0.833	-0.326	0.366	1.244	2.247	2.966	3.896	4.593	6.897
		NP	-1.931	-1.526	-1.363	-1.096	-0.833	-0.327	0.366	1.244	2.248	2.967	3.897	4.595	6.899
		HP	-1.933	-1.527	-1.364	-1.097	-0.834	-0.327	0.366	1.245	2.249	2.967	3.900	4.598	6.904

TABLA 3

ESTIMACION DEL SESGO DE PERCENTILES PARA MUESTRAS CON CORRELACION CERO

METODO	TAMANO DE MUESTRA	ESTIMACION DEL SESGO DE PERCENTILES PARA MUESTRAS CON CORRELACION CERO												
		F=0.001 K=-1.93	F=0.010 K=-1.53	F=0.020 K=-1.36	F=0.050 K=-1.10	F=0.100 K=-0.83	F=0.250 K=-0.33	F=0.500 K=0.37	F=0.750 K=1.25	F=0.900 K=2.25	F=0.950 K=2.97	F=0.980 K=3.90	F=0.990 K=6.91	
MO	5	0.472	0.395	0.365	0.315	0.266	0.171	0.041	-0.123	-0.311	-0.446	-0.521	-0.751	-1.183
MY		0.393	0.327	0.302	0.259	0.217	0.137	0.026	-0.114	-0.273	-0.388	-0.537	-0.648	-1.015
MP		0.008	0.007	0.007	0.006	0.005	0.004	0.002	0.000	-0.002	-0.004	-0.007	-0.008	-0.014
MO	10	0.240	0.201	0.166	0.160	0.136	0.089	0.023	-0.060	-0.154	-0.222	-0.310	-0.375	-0.592
MY		0.188	0.156	0.144	0.124	0.104	0.066	0.013	-0.053	-0.129	-0.184	-0.254	-0.307	-0.482
MP		-0.002	-0.001	-0.001	-0.001	0.000	0.002	0.002	0.004	0.007	0.008	0.010	0.012	0.017
MO	20	0.125	0.104	0.096	0.083	0.070	0.045	0.009	-0.034	-0.084	-0.120	-0.167	-0.201	-0.316
MY		0.094	0.078	0.072	0.061	0.051	0.032	0.004	-0.030	-0.069	-0.097	-0.134	-0.161	-0.251
MP		0.001	0.000	0.000	0.000	0.000	-0.000	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002	-0.003
MO	30	0.083	0.069	0.064	0.055	0.046	0.030	0.006	-0.023	-0.056	-0.080	-0.112	-0.135	-0.212
MY		0.061	0.050	0.046	0.039	0.033	0.020	0.002	-0.020	-0.045	-0.063	-0.087	-0.104	-0.163
MP		-0.001	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000	0.000	0.001	0.002
MO	50	0.049	0.041	0.038	0.032	0.027	0.018	0.004	-0.013	-0.032	-0.046	-0.064	-0.078	-0.122
MY		0.037	0.030	0.028	0.024	0.020	0.012	0.001	-0.012	-0.027	-0.038	-0.052	-0.062	-0.097
MP		-0.002	-0.002	-0.002	-0.002	-0.001	-0.000	-0.001	0.001	0.002	0.002	0.003	0.004	0.006
MO	80	0.035	0.029	0.027	0.023	0.019	0.013	0.002	-0.009	-0.023	-0.033	-0.046	-0.056	-0.087
MY		0.025	0.020	0.019	0.016	0.014	0.008	0.001	-0.008	-0.018	-0.025	-0.035	-0.042	-0.066
MP		0.002	0.002	0.002	0.001	0.001	0.001	0.000	-0.001	-0.001	-0.002	-0.003	-0.003	-0.005
MO	100	0.026	0.021	0.019	0.017	0.014	0.009	0.001	-0.007	-0.017	-0.025	-0.034	-0.041	-0.065
MY		0.018	0.014	0.013	0.011	0.009	0.006	0.000	-0.006	-0.013	-0.019	-0.026	-0.031	-0.048
MP		-0.001	-0.001	-0.001	-0.001	-0.001	-0.000	-0.001	0.000	-0.000	-0.000	0.000	0.001	0.001
MO	1000	0.003	0.002	0.002	0.002	0.001	0.001	0	-0.002	-0.003	-0.004	-0.006	-0.007	-0.010
MY		0.002	0.002	0.001	0.001	0.001	0	-0.001	-0.002	-0.003	-0.004	-0.005	-0.006	-0.008
MP		0	0	0	0	0	0	-0.001	-0.001	-0.001	-0.002	-0.002	-0.002	-0.003

TABLA 4

ESTIMACION DE LA VARIANZA DE PERCENTILES PARA MUESTRAS CON CORRELACION CERO

MÉTODO	TAMANO DE MUESTRA	F												
		0.001	0.010	0.020	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.980	0.990	0.999
NO	5	0.606	0.447	0.395	0.326	0.276	0.324	0.289	0.545	1.093	1.654	2.587	3.440	7.196
		0.520	0.369	0.347	0.292	0.254	0.226	0.285	0.523	1.015	1.513	2.338	3.069	6.383
		0.858	0.593	0.512	0.402	0.320	0.237	0.281	0.602	1.329	2.087	3.361	4.534	9.730
NO	10	0.384	0.261	0.227	0.181	0.147	0.115	0.140	0.287	0.612	0.948	1.511	2.028	4.311
		0.255	0.187	0.166	0.137	0.118	0.105	0.139	0.266	0.528	0.792	1.228	1.625	3.362
		0.373	0.263	0.228	0.179	0.144	0.110	0.138	0.296	0.644	1.003	1.605	2.147	4.594
NO	20	0.204	0.150	0.125	0.099	0.079	0.058	0.069	0.147	0.324	0.509	0.818	1.104	2.367
		0.123	0.090	0.080	0.066	0.057	0.051	0.069	0.149	0.313	0.484	0.770	1.032	2.187
		0.172	0.122	0.106	0.084	0.068	0.054	0.069	0.146	0.313	0.484	0.770	1.032	2.187
NO	30	0.142	0.100	0.086	0.068	0.054	0.039	0.047	0.101	0.224	0.352	0.568	0.767	1.648
		0.081	0.060	0.053	0.044	0.038	0.034	0.046	0.090	0.178	0.266	0.412	0.544	1.121
		0.112	0.079	0.069	0.055	0.044	0.036	0.046	0.098	0.210	0.323	0.513	0.687	1.451
NO	50	0.089	0.063	0.054	0.042	0.033	0.024	0.028	0.061	0.136	0.216	0.349	0.472	1.018
		0.048	0.035	0.031	0.026	0.022	0.020	0.028	0.057	0.106	0.159	0.246	0.325	0.670
		0.066	0.047	0.040	0.032	0.026	0.021	0.028	0.059	0.124	0.192	0.304	0.406	0.857
NO	80	0.058	0.040	0.035	0.030	0.021	0.015	0.017	0.038	0.087	0.138	0.224	0.303	0.656
		0.030	0.022	0.019	0.016	0.014	0.013	0.017	0.033	0.067	0.097	0.154	0.204	0.420
		0.041	0.029	0.025	0.020	0.016	0.013	0.017	0.036	0.078	0.120	0.190	0.254	0.536
NO	100	0.047	0.033	0.028	0.022	0.017	0.012	0.014	0.031	0.070	0.111	0.180	0.244	0.529
		0.024	0.017	0.015	0.013	0.011	0.010	0.014	0.027	0.053	0.080	0.123	0.163	0.335
		0.033	0.023	0.020	0.016	0.013	0.010	0.014	0.029	0.062	0.096	0.151	0.203	0.427
NO	1000	0.005	0.003	0.003	0.002	0.002	0.001	0.001	0.003	0.007	0.012	0.019	0.026	0.055
		0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.003	0.005	0.008	0.013	0.017	0.035
		0.003	0.002	0.002	0.002	0.001	0.001	0.001	0.003	0.006	0.010	0.015	0.021	0.043

ESTIMACION DEL ERROR CUADRATICO MEDIO DE PERCENTILES PARA MUESTRAS CON CORRELACION CERO

TAMANO DE MUESTRA	$F=0.001$	$F=0.010$	$F=0.020$	$F=0.050$	$F=0.100$	$F=0.250$	$F=0.500$	$F=0.750$	$F=0.900$	$F=0.950$	$F=0.980$	$F=0.990$	$F=0.999$
5	0.828	0.603	0.528	0.425	0.347	0.353	0.291	0.560	1.190	1.853	2.858	4.004	8.596
MP	0.674	0.486	0.436	0.359	0.301	0.245	0.286	0.536	1.090	1.664	2.626	3.509	7.414
NP	0.838	0.593	0.512	0.402	0.320	0.237	0.281	0.602	1.329	2.087	3.361	4.534	9.730
10	0.421	0.301	0.262	0.207	0.166	0.123	0.141	0.291	0.636	0.997	1.607	2.169	4.662
MP	0.290	0.211	0.187	0.152	0.129	0.109	0.139	0.269	0.545	0.826	1.292	1.719	3.595
NP	0.373	0.263	0.228	0.179	0.144	0.110	0.138	0.296	0.644	1.003	1.605	2.147	4.594
20	0.220	0.161	0.134	0.106	0.084	0.060	0.069	0.148	0.331	0.523	0.846	1.144	2.467
MP	0.132	0.086	0.065	0.070	0.060	0.052	0.069	0.150	0.318	0.493	0.768	1.058	2.250
NP	0.172	0.122	0.106	0.084	0.068	0.054	0.069	0.146	0.313	0.484	0.770	1.032	2.187
30	0.149	0.105	0.090	0.071	0.056	0.040	0.047	0.102	0.227	0.358	0.581	0.785	1.693
MP	0.085	0.063	0.055	0.046	0.039	0.034	0.046	0.090	0.180	0.270	0.420	0.555	1.148
NP	0.112	0.079	0.069	0.055	0.044	0.036	0.046	0.098	0.210	0.323	0.513	0.687	1.451
50	0.091	0.065	0.055	0.043	0.034	0.024	0.028	0.061	0.137	0.218	0.353	0.478	1.033
MP	0.049	0.036	0.032	0.027	0.022	0.020	0.026	0.057	0.107	0.160	0.249	0.329	0.679
NP	0.066	0.047	0.040	0.032	0.026	0.021	0.028	0.059	0.124	0.192	0.304	0.406	0.857
80	0.059	0.041	0.036	0.031	0.021	0.015	0.038	0.087	0.015	0.225	0.305	0.359	0.808
MP	0.031	0.022	0.019	0.016	0.013	0.017	0.033	0.067	0.097	0.155	0.205	0.422	0.904
NP	0.041	0.029	0.025	0.020	0.016	0.013	0.017	0.036	0.078	0.120	0.190	0.254	0.536
100	0.048	0.033	0.028	0.022	0.017	0.012	0.014	0.031	0.070	0.112	0.151	0.246	0.533
MP	0.024	0.017	0.015	0.013	0.011	0.010	0.014	0.027	0.053	0.080	0.124	0.164	0.337
NP	0.033	0.023	0.020	0.016	0.013	0.010	0.014	0.029	0.062	0.096	0.151	0.203	0.427
1000	0.005	0.003	0.003	0.002	0.002	0.001	0.001	0.003	0.007	0.012	0.019	0.026	0.055
MP	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.003	0.005	0.008	0.013	0.017	0.035
NP	0.003	0.002	0.002	0.002	0.001	0.001	0.001	0.003	0.006	0.010	0.015	0.021	0.043

TABLA 5

COMPARACION DEL SESGO POR LOS 3 METODOS
EN MUESTRAS ALEATORIAS PARA AFTA

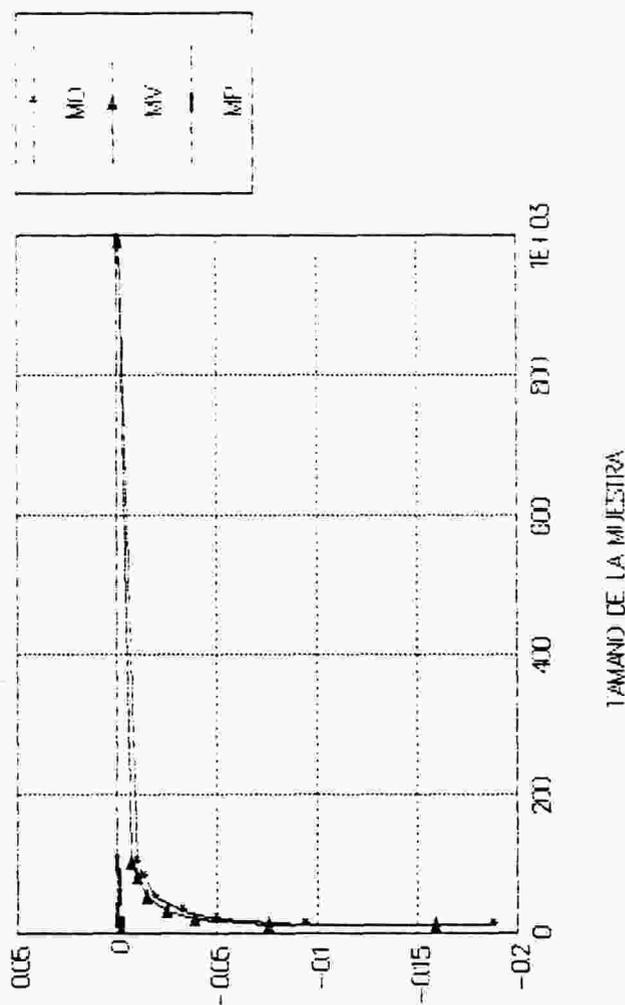
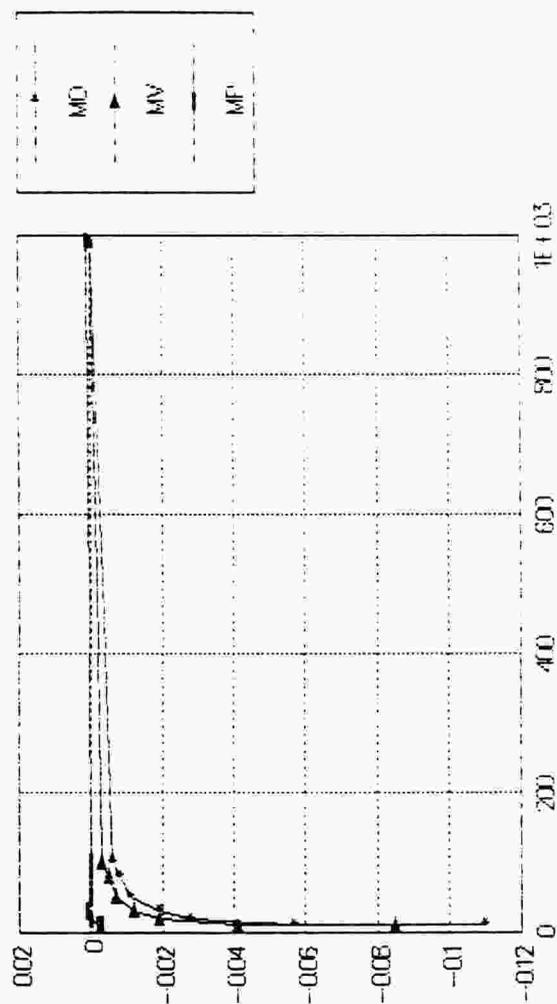


Figura 1

COMPARACION DEL SESGO POR LOS 3 METODOS

EN MUESTRAS ALATORIAS PARA MIJ



TAMANO DE LA MUESTRA

00535

Figura 2

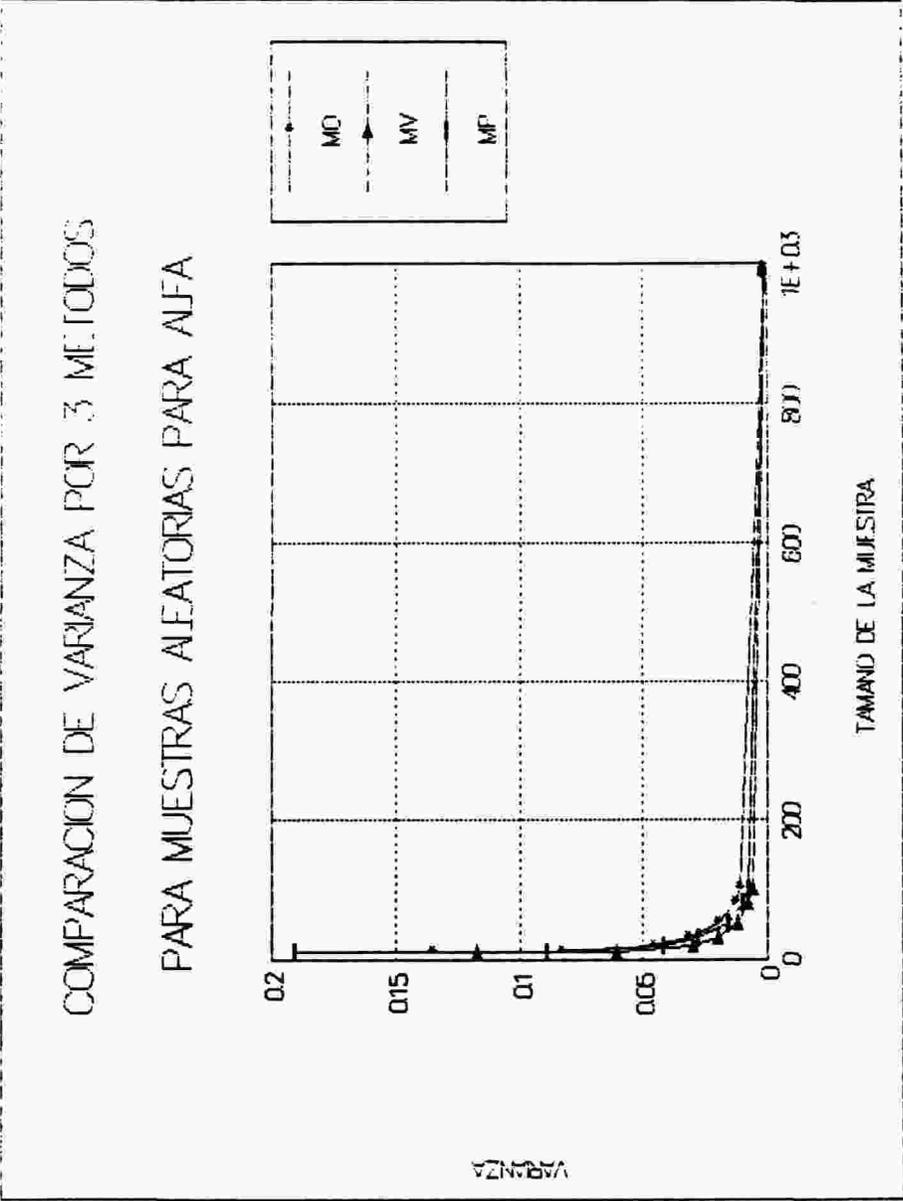


Figura 3

COMPARACION DE ECM POR 3 METODOS
 EN MUESTRAS ALEATORIAS PARA ALFA

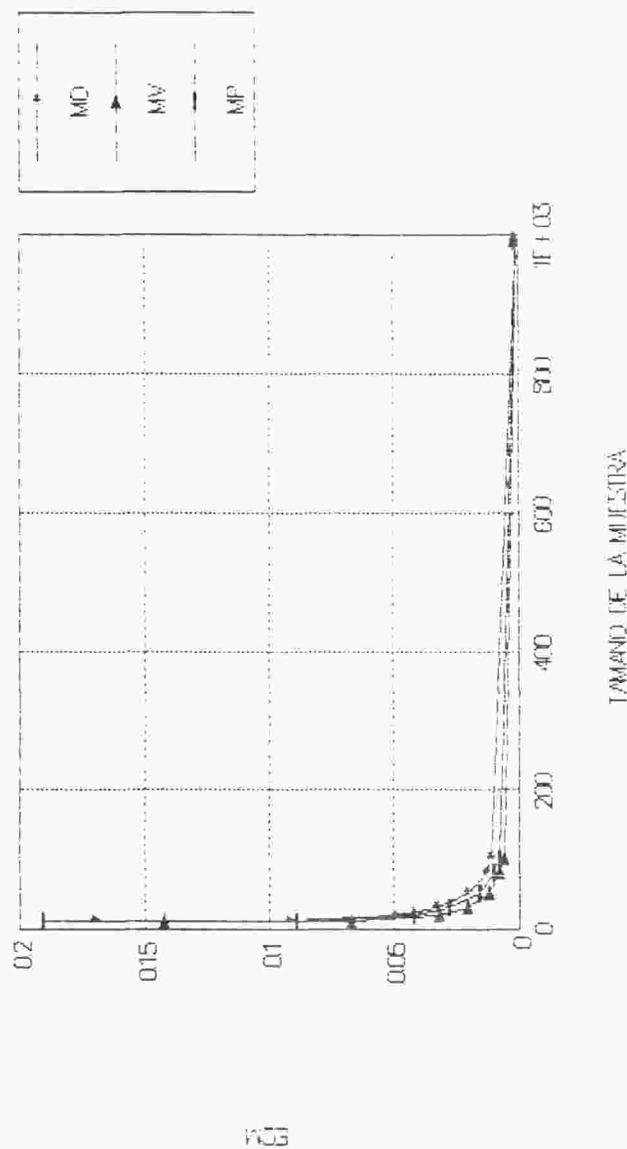


Figura 4

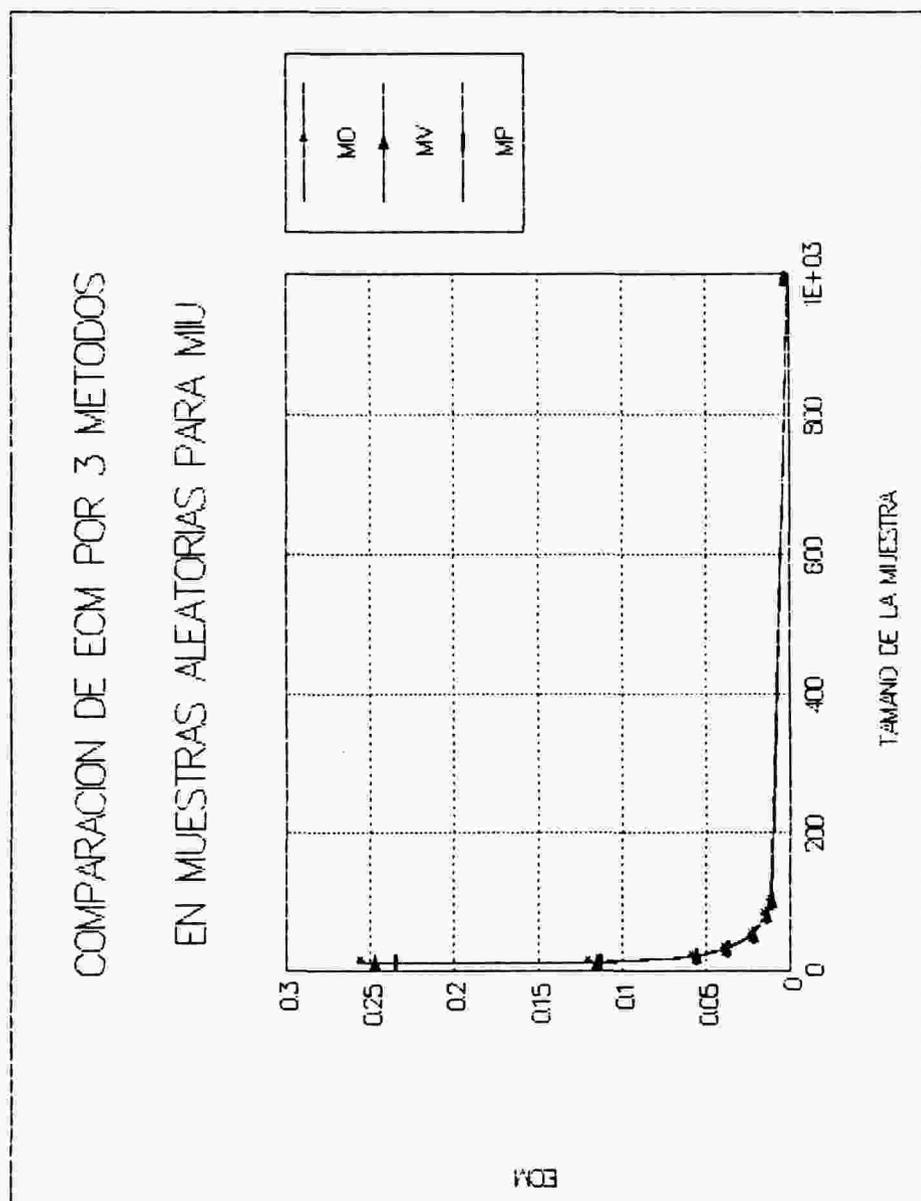


Figura 5

BIBLIOGRAFIA

- Ahrens, J.H., U. Dieter (1970), Pseudo-random Numbers: A new proposal for choice of Multipliers. *Computing* 12, 223-246.
- Allard, J.L., A.R. Dobell and T.E. Hull (1963). Mixed, Congruential Random Number Generators for Decimal Machine. *JACM* 10, 131-141.
- Arvillias, A.C., and D.G. Maritsas (1978). Particioning the period of a Class m -sequences and Applications to pseudorandom Number Generation, *JACM* 25, 675-686.
- Barnett, V.D. (1962), The Behavior of pseudo-random sequence Generated on Computers by the Multiplicative Congruential Method. *Math.* 16, 63-69.
- Bedall, F.K. and H. Zimmermann (1976). On the Generation of $N(\mu, \alpha)$ -Distributed Random Vectors by $N(0, 1)$ -distributed Random Numbers. *Biometrische Zeitschrift* 18, 467-472.
- Bell J.R. (1968). Normal Random Deviates, *CACM* 11, 498.
- Butler, E.L. (1970). Algorithm 270: General Random Number Generator, *CACM* 13, 49-52.
- Clay, S.C.R.D. Fardo and M. Manzumdar (1975). On using the Box-Muller Transformation with Multiplicative Congruential Pseudo-Random Number Generators, *Appl. Stat.* 24, 132-135.
- Claustriax, J.J. (1976). Generation and validity Control of Pseudo-Random Number on a Computer with 16 Bit Words, *Revue de Statistique Appliques* 24, 75-88.
- Coldwell, R.L. (1974). Correlational Defects in the Standart IBM 360 Random Number Generator and Clasical Ideal

Gas Correlational Function. *J. Computational physics* 14, 223-226.

- Dieter, U. (1972). Statistical Interdependence of Pseudo Random Number Generated by the linear Congruential Method, in *Applications of Number Theory to Numerical Analysis*. edited by S.K. Zeremba. Academic Press. New York. 287-318.
- Downham, D.Y. and F.D.K. Robers (1967). Multiplicative Congruential Pseudo-Random Number Generators. *Comp. J.* 10, 74-77.
- Fuller A.T. (1976). The Period of Pseudo-Random Numbers Generated by Lehmer's Congruential Method. *Comp. J.* 19. 173-177.
- Greenberger, M. (1961a). Notes on a New Pseudo-Random Number generator, *JACM* 8, 163-167.
- Greenwood, J.A. Landwehr, J.M. Matalas, N.C., and Wallis, J. R. (1979). Probability Weighted Moments: Definition and Relation to parameters of several Distributions Expressable in Inverse form.
- Galombos, J. (1978). *The Asymptotic theory of extreme Order Statistics*. Wiley, New York.
- Hemmerle, W.J. (1969). Generating Pseudo-Random Numbers on a Two's Complement Machine such as the IBM 360, *CACM* 12, 382-383.
- Hosting, J.R.M. (1984). Testing Whether the shape parameter is Zero in the generalized extreme - value distribution. *Biometrika* 71, in press.
- Johnson, M.E. and J.S. Ramberg (1977). Elliptically Symmetric Distributions: Characterizations and Random Variate Generation. *Proceedings of the Statistical Computing Section. American Statistical Association, Wash-*

- ington, D.C., 262-265.
- Jenkinson, A.T.F. (1969). Statistics of extremes WMO Technical NOTE 98. Geneva: World Meteorological office.
- Lewis, P.A.W., O.S. Goodman, and J.W. Miller (1969). A Pseudo-random Number Generator for the System 360, IBM. Systems J. 8, 136-146.
- Liniger, W. (1961). On a Method by D.H. Lehmer for the generation of Pseudo-Random Numbers. Numer. Math. 3. 265-270.
- Landwehr J.M., Matalas, N.C. and Wallis. J.R., Some Comparisons of flood statistics in real and log Space, Water Resour. res., 14 (5) 902-920, 1978.
- Landwehr, J.M.: Matalas, N.C. and Wallis, J.R. (1979). Probability Weighted moments compared with some traditional Techniques in estimating Gumbell parameter and quantiles. Water Resour. rec: 15.
- Mars, P. and A.J. Miller (1977). Theory and Design of a Digital Stochastic Computer Random Number Generator, Mathematics and Computers in Simulation 19, 193-216.
- Smith W.B., and R.R. Hocking (1972). Wishart Variate Generator, Appl. Stat. 21. 341-345.