

CONTRASTE DE MODELOS PROBABILISTICOS DESDE UNA PERSPECTIVA BAYESIANA

José-Miguel Bernardo
Departamento de Bioestadística
Universidad de Valencia

SUMMARY

It is pointed out that the problem of testing whether the data are compatible or not with a fully specified probabilistic model may be reduced to the problem of testing for uniformity a suitable transformation of the data. By embedding the uniform distribution into a new family of distributions, easily tractable, a Bayesian test for uniformity is presented, which therefore constitutes an alternative to classical model checking.

Key Words: Bayesian Inference; Bayesian Testing; Model Checking; Reference Distributions.

RESUMEN

Se pone de manifiesto que el problema de contrastar si un conjunto de datos es o no compatible con un modelo probabilístico totalmente especificado puede ser reducido al problema de contrastar si una determinada transformación de los datos puede ser considerada como una muestra aleatoria de una distribución uniforme. Mediante la construcción de una nueva familia paramétrica de distribuciones, que contiene a la uniforme como caso particular, se propone un contraste bayesiano de uniformidad que constituye una alternativa a los métodos clásicos para el contraste de modelos probabilísticos.

Palabras Clave: Contraste Bayesiano; Distribuciones de Referencia; Inferencia Bayesiana; Verificación de Modelos.

Clasificación AMS (1980).

Primaria: 62F03, 62H15.

Secundaria: 62F15, 62A15, 62E25.

1. INTRODUCCION

Sea $\{x_1, x_2, \dots, x_n\}$ un conjunto de n observaciones unidimensionales y supongamos que desea contrastarse la hipótesis de que constituyen una muestra aleatoria de una distribución cuya función de distribución es F_0 ; podría tratarse, por ejemplo, de contrastar si las x_i pueden ser consideradas como observaciones independientes de una *determinada* distribución normal.

Si la hipótesis es correcta, es inmediato comprobar (ver e.g. Renyi, 1976 p. 204) que la imagen $\{y_1, y_2, \dots, y_n\}$ del conjunto de datos $\{x_1, x_2, \dots, x_n\}$ mediante la aplicación $F_0: R \rightarrow [0, 1]$ definida por la función de distribución propuesta, debe ser una muestra aleatoria de una distribución uniforme en $[0, 1]$. Consecuentemente, para probar la incompatibilidad de las x_i con la distribución supuesta bastaría probar la incompatibilidad de las y_i con la distribución uniforme.

En las Secciones 2, 3 y 4 desarrollaremos un procedimiento que permite contrastar la hipótesis de uniformidad. La idea básica consiste en construir una familia paramétrica de distribuciones de probabilidad que sea matemáticamente tratable y que contenga a la distribución uniforme, y a un conjunto de distribuciones de probabilidad razonablemente denso en un entorno suyo, como casos particulares; tomando esta familia como modelo probabilístico de las y_i , se procede entonces a contrastar la hipótesis de que el verdadero valor del parámetro introducido es el que corresponde a la distribución uniforme. Si esta hipótesis puede ser considerada incompatible con las y_i , entonces la distribución F_0 puede ser considerada incompatible con las x_i . Naturalmente, el hecho de que los datos no permitan rechazar F_0 *no prueba* que F_0 sea cierta; tan solo permite mantener provisionalmente F_0 como una *posible* explicación de los datos obtenidos.

En la Sección 5 se ofrece un ejemplo numérico con el que se ilustra el comportamiento del procedimiento descrito y finalmente, en la Sección 6, se discuten los resultados obtenidos y se sugieren nuevas líneas de investigación.

2. LA FAMILIA DE DISTRIBUCIONES ALFA

Ya hemos mencionado que buscamos una familia paramétrica de distribuciones que contenga como casos particulares a la distribución uniforme y a un conjunto razonablemente denso de distribuciones de probabilidad que puedan considerarse cercanas a ella. La familia de distribuciones Beta parece un candidato obvio pero, desgraciadamente, resulta intratable para nuestros propósitos; en consecuencia, hemos construido una nueva familia de distribuciones sobre $[0, 1]$, emparentada con la familia Beta pero mucho más manejable como modelo probabilístico, a la que hemos denominado familia de distribuciones Alfa.

DEFINICION 1. Una cantidad aleatoria absolutamente continua Y tiene una *distribución Alfa* si su función de densidad de probabilidad, que denotaremos $Al(y|\theta, \omega)$ es de la forma:

$$\begin{aligned} p(y|\theta, \omega) = Al(y|\theta, \omega) &= \theta \{y/\omega\}^{\theta-1}, \quad 0 \leq y \leq \omega \\ &= \theta \{(1-y)/(1-\omega)\}^{\theta-1}, \quad \omega \leq y \leq 1 \end{aligned}$$

y se anula fuera del intervalo unidad, y donde $\theta > 0$ y $0 < \omega < 1$.

Puede comprobarse que la familia de distribuciones así definida se reduce a la distribución uniforme para $\theta = 1$, *cualquiera* que sea el valor de ω . Integrando las expresiones anteriores se obtiene inmediatamente la correspondiente función de distribución, que resulta ser:

$$\begin{aligned} F_Y(y|\theta, \omega) &= \omega \{y/\omega\}^{\theta}, \quad 0 \leq y \leq \omega \\ &= 1 - (1-\omega) \{(1-y)/(1-\omega)\}^{\theta}, \quad \omega \leq y \leq 1 \end{aligned}$$

En la Figura 1 se ofrece la representación gráfica de las funciones de distribución y de densidad de probabilidad de algunas distribuciones Alfa.

La función de verosimilitud correspondiente a una muestra aleatoria $z = \{y_1, \dots, y_n\}$ de una distribución $Al(y|\theta, \omega)$ resulta ser

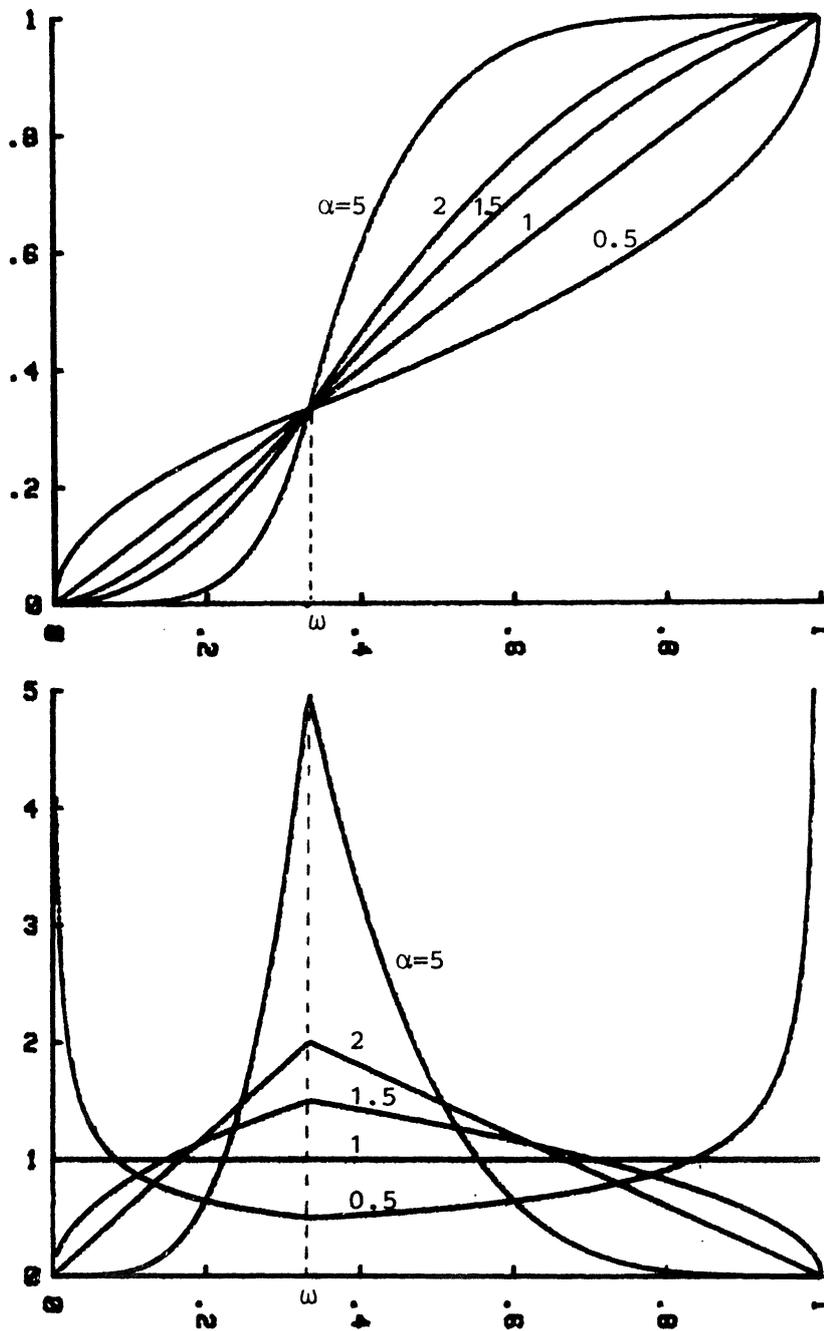


Fig. 1. Funciones de distribución y funciones de densidad de probabilidad de las distribuciones Alfa con parámetros $\omega = 1/3$ y $\alpha = 0.5, 1, 1.5, 2$ y 5 .

$$\begin{aligned}
p(z|\theta, \omega) &= \prod_{i=1}^n p(y_i|\theta, \omega) = \\
&= \theta^n \left\{ \prod_{i \in I(\omega)} \left(\frac{y_i}{\omega} \right) \prod_{j \in J(\omega)} \left(\frac{1-y_j}{1-\omega} \right) \right\}^{\theta-1}
\end{aligned}$$

donde $I(\omega)$ y $J(\omega)$ son, respectivamente, los conjuntos de índices correspondientes a los conjuntos de observaciones $\{y_i; y_i \leq \omega\}$ y $\{y_j; y_j > \omega\}$ en los que el valor de ω divide a los elementos de la muestra. Definiendo

$$t(\omega) = \frac{1}{n} \left\{ \sum_{i \in I(\omega)} \log \left(\frac{\omega}{y_i} \right) + \sum_{j \in J(\omega)} \log \left(\frac{1-\omega}{1-y_j} \right) \right\} \quad (1)$$

la función de verosimilitud $p(z|\theta, \omega)$ puede ser expresada en la forma

$$p(z|\theta, \omega) = \theta^n \exp \{-n(\theta-1)t(\omega)\} \quad (2)$$

Puesto que toda la información proporcionada por la muestra se encuentra contenida en la *función* $t(\omega)$, resulta conveniente estudiar su comportamiento. Denotaremos por $\text{Ga}(x|\alpha, \beta)$ a la densidad de probabilidad Gamma, de forma que $\text{Ga}(x|\alpha, \beta) = \beta^\alpha x^{\alpha-1} e^{-\beta x} / \Gamma(\alpha)$.

La densidad de probabilidad de una observación y de $\text{Al}(y|\theta, \omega)$ condicionada a $y \leq \omega$ resulta ser

$$p(y|\theta, \omega, y \leq \omega) = \text{Al}(y|\theta, \omega) / p\{y \leq \omega\} = (\theta/\omega) (y/\omega)^{\theta-1}$$

y por tanto, si $s = \log y$, $y = e^s$

$$p(s|\theta, y \leq \omega) = p(y|\theta, y \leq \omega) / |ds/dy| = \theta e^{s\theta} / \omega^\theta$$

Consecuentemente, si $t_1 = \log y - \log \omega$

$$p(t_1|\theta, y \leq \omega) = \theta e^{t_1\theta} = \text{Ga}(t_1|1, \theta)$$

Análogamente, si $t_2 = \log(1 - y) - \log(1 - \omega)$ resulta, de nuevo, $p(t_2 | \theta, y > \omega) = \text{Ga}(t_2 | 1, \theta)$. En consecuencia, a la vista de (1), $t(\omega) = \sum t_i / n$ y, por tanto

$$p\{t(\omega) | \theta\} = \text{Ga}\{t(\omega) | n, n\theta\}$$

Tenemos pues, que para el verdadero y desconocido valor de ω , la distribución de $t(\omega)$ es una Gamma con parámetros n y $n\theta$; consecuentemente, resulta $E\{t(\omega) | \theta\} = 1/\theta$ y $D^2\{t(\omega) | \theta\} = 1/n\theta^2$ y por tanto, dado ω , $1/t(\omega)$ es un estimador consistente de θ .

3. INFERENCIA SOBRE LOS PARAMETROS DE UNA DISTRIBUCION ALFA

Puesto que la distribución uniforme corresponde a $\theta = 1$ para cualquier valor de ω , nuestro objetivo es estimar el valor de θ mediante la muestra $\{y_1, \dots, y_n\}$ y comprobar si es o no es significativamente distinto de 1. Desde un punto de vista Bayesiano el problema se reduce a determinar la distribución final del parámetro de interés θ eliminando por integración el parámetro marginal ω .

La distribución final del parámetro de interés dependerá obviamente de la distribución inicial que se utilice. Si no se dispone de mas información sobre los valores de θ y ω que la proporcionada por los datos, o si no se quiere utilizar información subjetiva, debe utilizarse la correspondiente distribución mínimo informativa o de referencia, que puede calcularse mediante el método de maximización de la información desconocida (Bernardo, 1979), según el cual

$$\pi(\omega | \theta) \propto \lim_{n \rightarrow \infty} \exp\{E_{z|\theta, \omega} [\log p(\omega | \theta, z)]\} \quad (3)$$

$$\pi(\theta) \propto \lim_{n \rightarrow \infty} \exp\{E_{z|\theta} [\log p(\theta | z)]\} \quad (4)$$

utilizando (2), tenemos obviamente que, para n grande,

$$p(\omega | \theta, z) \propto \exp\{-n(\theta - 1)t(\omega)\} \propto \exp\{-n(\theta - 1)/\hat{\theta}\}$$

puesto que $t(\omega)$ es, según hemos visto, un estimador consistente de θ . Consecuentemente, utilizando (3) resulta $\pi(\omega|\theta) \propto \exp\{-n(\theta - 1)/\theta\}$, independiente de ω y por lo tanto, al normalizar, resulta $\pi(\omega|\theta) \propto 1$.

Análogamente, utilizando de nuevo (2), para n grande,

$$p(\theta|z) \propto \theta^n \exp\{-n(\theta - 1)/\hat{\theta}\} \propto \text{Ga}(\theta|n + 1, n/\hat{\theta})$$

y, por lo tanto, haciendo uso de (4)

$$\pi(\theta) \propto \exp\{E_{z|\theta}[\log \text{Ga}(\theta|n + 1, n/\hat{\theta})]\} \propto \theta^{-1}$$

Consecuentemente, la distribución inicial de referencia para hacer inferencias sobre el parámetro de interés θ resulta ser

$$\pi(\theta) \pi(\omega|\theta) \propto \theta^{-1}$$

con lo que, utilizando (2), la distribución final conjunta de referencia será:

$$\begin{aligned} \pi(\theta, \omega|z) &\propto p(z|\theta, \omega) \pi(\omega|\theta) \pi(\theta) \\ &\propto \theta^{n-1} \exp\{-n(\theta - 1)t(\omega)\} \end{aligned}$$

que puede expresarse en la forma

$$\pi(\theta, \omega|z) \propto \text{Ga}(\theta|n, nt(\omega)) e^{nt(\omega)}/t(\omega)^n$$

En consecuencia, podemos escribir

$$\pi(\theta|\omega, z) = \text{Ga}(\theta|n, nt(\omega)) \quad (5)$$

$$\pi(\omega|z) \propto e^{nt(\omega)}/t(\omega)^n \quad (6)$$

y, finalmente

$$\pi(\theta|z) = \int \text{Ga}(\theta|n, nt(\omega)) \pi(\omega|z) d\omega \quad (7)$$

$$\propto \theta^{n-1} \int_0^1 \exp\{-n(\theta - 1)t(\omega)\} d\omega \quad (8)$$

El cálculo del resultado exacto (8) es relativamente elemental si se dispone de un ordenador. La expresión (5) permite no obstante obtener una aproximación asintótica que resulta eficaz para tamaños muestrales no muy pequeños. En efecto, para cualquier estimador consistente $\hat{\omega}$ del parámetro marginal, y n suficientemente grande,

$$\pi(\theta|z, \hat{\omega}) \simeq \pi(\theta|z)$$

de forma que

$$\pi^*(\theta|z) \simeq \text{Ga} \{ \theta | n, n t(\hat{\omega}) \} \quad (9)$$

es una aproximación asintótica a la distribución final del parámetro de interés. Puesto que la función de distribución Alfa verifica que $F(\omega) = \omega$ para todo θ , es fácil obtener gráficamente un estimador de ω como el punto en el que la función de distribución empírica de la muestra $\{y_1, \dots, y_n\}$ corta a la recta $F(y) = y$. Alternativamente, puesto que (6) es una función creciente en $t(\omega)$ para $t(\omega) > 1$ y decreciente para $t(\omega) < 1$, la moda de la distribución final $\pi(\omega|z)$ coincide necesariamente con un extremo relativo de la función $t(\omega)$, lo que proporciona un sencillo procedimiento analítico para obtener un buen estimador de ω .

4. CONTRASTE DE LA HIPOTESIS $\theta = 1$

Una vez obtenida la distribución final del parámetro de interés (7), o una aproximación suya como (9), debe decidirse si el resultado es o no compatible con la hipótesis $\theta = 1$. A nivel intuitivo, basta con observar si el valor $\theta = 1$ pertenece o no a una región de alta densidad de la distribución final de θ ; analíticamente, puede determinarse si el punto $\theta = 1$ pertenece o no a un intervalo creíble o de confianza de un nivel preestablecido.

Partiendo de la distribución final aproximada (9) y utilizando el hecho conocido (Basulto, 1978) de que si $p(\theta) = \text{Ga}(\theta|\alpha, \beta)$ entonces $\xi = \log \theta$ tiene una distribución *aproximadamente* normal, $p(\xi) \sim N(\xi|\mu, \sigma)$ con

$$\mu = \log(\alpha/\beta) - 1/(2\alpha); \quad \sigma = 1/\sqrt{\alpha}$$

el contraste de $\theta = 1$ cuando $\pi(\theta|z) = \text{Ga}\{\theta|n, n t(\hat{\omega})\}$ puede reducirse, aproximadamente, al de $\xi = 0$ cuando

$$\pi(\xi|z) = N(\xi| \{-\log t(\hat{\omega}) - 1/(2n)\}, 1/\sqrt{n})$$

para lo que basta calcular

$$\gamma = \frac{|\mu|}{\sigma} = |\sqrt{n} \log t(\hat{\omega}) + 1/2\sqrt{n}| \quad (10)$$

y comparar su valor con los que proporcionan las tablas de la distribución normal.

Valores grandes de γ , en la práctica valores mayores que 3, indican la *incompatibilidad* de los datos $\{x_1, \dots, x_n\}$ con la distribución propuesta F_0 ; valores pequeños de γ indican la compatibilidad de los datos con tal hipótesis, lo que, naturalmente, *no prueba* que la hipótesis sea correcta: simplemente no la elimina.

Bernardo (1980) analiza desde un punto de vista Bayesiano este problema clásico de contraste de hipótesis, encontrando una relación biyectiva entre los niveles de significación clásicos y la probabilidad final asociada a la hipótesis nula. Específicamente, la probabilidad asociada a la hipótesis contrastada resulta ser la función monótona del estadístico γ

$$p(H_0|\gamma) = 1/[1 + \exp\{\frac{1}{2}(\gamma^2 - 1)\}] \quad (11)$$

Así, a un resultado significativo al 5% desde un punto de vista clásico ($\gamma = 1.96$) le corresponde una probabilidad final de 0.195, todavía bastante apreciable, mientras que para $\gamma = 3.19$ (nivel de significación aproximado 0.1%) le corresponde una probabilidad final de 0.01.

5. EJEMPLO NUMERICO

Considerense las muestras I y II, de tamaño 250, reproducidas en la Tabla 1 y obtenidas respectivamente por simulación de una distribución normal $N(x | 2, \sqrt{2})$ y de una distribución Gamma, $Ga(x | 2, 1)$. Para comprobar que el procedimiento propuesto permite discriminar entre *formas* alternativas, las distribuciones se han elegido con los dos primeros momentos iguales: $E(x) = 2$ y $D^2(x) = 2$; consecuentemente, la distribución normal escogida, es la *mejor* aproximación normal posible a la distribución $Ga(x | 2, 1)$, como se observa en la Figura 2.

Se ha contrastado en ambos casos la hipótesis de que se trata de observaciones independientes de una distribución $N(x | 2, \sqrt{2})$. Los datos transformados mediante la función de distribución correspondiente a esta distribución normal son los que se reproducen en la Tabla 2.

Las distribuciones finales del parámetro marginal ω correspondientes a cada una de las muestras, aparecen en la Figura 3. Puede observarse que, como cabía esperar, la distribución correspondiente a la muestra I, que es verdaderamente normal $N(x | 2, \sqrt{2})$, es prácticamente uniforme, mientras que la correspondiente a la muestra II se concentra hacia la derecha con una moda en 0.92.

Las distribuciones finales del parámetro de interés θ , obtenidas por integración numérica a partir de la ecuación (7) son las representadas en la Figura 4. Como puede observarse, la distribución correspondiente a la primera muestra contiene claramente al punto $\theta = 1$, con lo que se mantiene la hipótesis de normalidad, mientras que la distribución correspondiente a la segunda muestra se sitúa completamente a su izquierda, rechazando así la hipótesis de normalidad.

Aunque la conclusión resulta obvia a la vista de las representaciones gráficas, puede obtenerse numéricamente el nivel de confianza necesario en cada caso para incluir al punto $\theta = 1$ en un intervalo de confianza. Estos niveles resultan ser, respectivamente, 0.126 y 0.99996: desde un punto de vista clásico, la hipótesis de normalidad sería rechazada en el caso de la muestra II para todo nivel de significación mayor que 0.00004; desde un punto de vista Bayesiano la probabilidad asociada a la hipótesis nula en este mismo caso es 0.0005.

Tabla 1. Muestras de $N(x|2, \sqrt{2})$ y de $Ga(x|2, 1)$ de 250 elementos

MUESTRA I

1.347	2.818	1.690	1.765	3.143	3.709	2.292	1.395	.170	1.133
5.727	1.858	2.969	2.693	3.017	2.081	3.027	1.817	.652	.547
3.580	6.130	2.104	.424	1.930	3.380	2.478	-1.185	1.017	2.812
1.084	1.243	-.004	3.253	-.130	2.255	3.027	1.272	-.482	1.541
1.393	2.294	3.553	2.344	2.053	5.243	3.192	.734	3.092	5.464
4.928	3.307	2.367	3.127	.792	.710	-.504	1.001	2.770	2.448
.720	2.691	1.249	.773	.666	1.332	2.321	.285	3.070	.358
.991	3.375	3.653	.055	.277	3.151	.975	-.317	2.617	2.012
1.601	3.491	1.306	.316	1.410	1.694	1.380	3.502	3.407	2.068
2.435	.856	3.899	2.050	1.808	4.152	.273	1.508	-.248	.219
1.783	-2.390	3.204	-.417	3.238	2.863	1.225	1.714	4.123	3.698
1.276	3.726	1.970	1.235	1.314	2.863	1.513	5.560	-.940	3.376
4.709	2.659	1.896	3.217	.693	1.732	.891	-.991	2.007	.933
1.822	.987	1.080	2.183	3.675	2.596	5.727	1.860	.889	2.735
1.924	1.364	1.705	4.175	3.364	1.065	2.571	1.836	2.345	2.635
1.341	2.056	3.621	.206	-1.346	2.256	4.490	-.085	-.096	.661
2.029	2.032	.899	1.497	2.236	1.785	2.005	2.895	1.936	1.637
1.207	3.040	3.094	1.292	.605	.455	1.722	4.050	-.018	2.502
1.456	1.757	5.411	1.333	3.111	1.142	2.296	1.801	1.665	2.233
2.721	1.787	-1.396	3.645	-.910	2.615	4.440	1.928	1.582	2.449
1.206	1.544	5.316	5.052	5.213	1.136	2.076	2.536	3.457	2.529
1.297	.539	2.909	1.379	2.261	.923	2.763	.628	2.043	3.260
.642	2.599	2.037	1.174	-.958	1.435	3.800	2.374	.973	2.515
1.037	2.034	2.204	3.313	2.086	3.756	3.130	2.750	2.485	1.993
3.672	1.154	.412	-2.125	1.961	1.414	.425	2.540	3.149	3.277

MUESTRA II

1.124	2.345	1.334	1.385	2.758	3.661	1.802	1.151	.624	1.010
10.039	1.451	2.529	2.203	2.590	1.622	2.603	1.422	.794	.753
3.432	12.284	1.640	.708	1.504	3.106	1.979	.317	.953	2.338
.985	1.067	.572	2.914	.537	1.769	2.602	1.082	.450	1.238
1.150	1.804	3.387	1.850	1.599	7.881	2.827	.827	2.689	8.805
6.733	2.994	1.872	2.737	.851	.817	.445	.945	2.289	1.949
.821	2.201	1.070	.844	.800	1.115	1.829	.661	2.659	.685
.940	3.098	3.560	.589	.658	2.769	.933	.489	2.121	1.567
1.276	3.283	1.101	.671	1.160	1.337	1.142	3.301	3.147	1.611
1.936	.879	4.025	1.597	1.415	4.568	.657	1.218	.506	.639
1.397	.173	2.843	.465	2.893	2.398	1.057	1.350	4.502	3.640
1.085	3.692	1.534	1.062	1.105	2.398	1.221	9.235	.358	3.099
6.036	2.165	1.479	2.862	.810	1.362	.895	.349	1.563	.914
1.425	.939	.983	1.707	3.599	2.099	10.040	1.452	.894	2.250
1.499	1.133	1.344	4.622	3.081	.976	2.072	1.435	1.851	2.139
1.120	1.602	3.503	.635	.292	1.770	5.410	.549	.546	.797
1.581	1.583	.898	1.211	1.753	1.399	1.562	2.436	1.509	1.299
1.048	2.620	2.691	1.093	.775	.719	1.456	4.341	.568	2.002
1.187	1.379	8.573	1.116	2.714	1.014	1.806	1.410	1.317	1.750
2.233	1.400	.285	3.545	.363	2.118	5.277	1.503	1.264	1.949
1.047	1.240	8.175	7.165	7.763	1.011	1.618	2.036	3.227	2.029
1.096	.750	2.454	1.142	1.775	.909	2.281	.784	1.592	2.924
.790	2.101	1.587	1.031	.355	1.175	3.832	1.877	.932	2.015
.963	1.584	1.725	3.003	1.626	1.747	2.741	2.267	1.985	1.552
3.593	1.021	.704	.198	1.528	1.162	.709	2.041	2.766	2.949

Tabla 2. Imágenes de las muestras I y II mediante la función de distribución de la normal $N(x|2, \sqrt{2})$

IMAGEN MUESTRA I

.322	.719	.413	.434	.791	.887	.582	.334	.098	.270
.996	.460	.753	.688	.764	.523	.766	.448	.170	.152
.868	.998	.529	.133	.480	.835	.633	.012	.243	.717
.258	.296	.078	.812	.066	.572	.766	.303	.040	.373
.334	.582	.864	.596	.515	.989	.800	.185	.780	.993
.981	.822	.603	.787	.196	.181	.038	.240	.707	.624
.183	.688	.297	.193	.173	.318	.590	.113	.775	.123
.238	.835	.879	.085	.112	.792	.234	.051	.669	.503
.389	.854	.312	.117	.338	.414	.330	.856	.840	.511
.621	.209	.910	.514	.445	.936	.111	.364	.056	.104
.439	.001	.803	.044	.809	.729	.292	.420	.933	.885
.304	.889	.491	.294	.314	.729	.365	.994	.019	.835
.972	.680	.471	.805	.178	.425	.216	.017	.502	.225
.450	.237	.257	.552	.882	.663	.996	.460	.216	.699
.478	.326	.417	.938	.833	.254	.657	.454	.597	.675
.320	.516	.874	.102	.009	.572	.961	.070	.069	.172
.508	.509	.218	.361	.567	.439	.501	.737	.482	.399
.287	.769	.780	.308	.162	.137	.422	.926	.077	.639
.350	.432	.992	.318	.784	.272	.583	.444	.406	.566
.695	.440	.008	.878	.020	.668	.958	.480	.384	.625
.287	.373	.990	.985	.988	.271	.522	.648	.849	.646
.309	.151	.740	.330	.573	.223	.705	.166	.512	.814
.168	.664	.510	.279	.018	.345	.898	.604	.234	.642
.248	.510	.557	.823	.524	.893	.788	.702	.634	.498
.881	.275	.131	.002	.489	.339	.133	.649	.792	.817

IMAGEN MUESTRA II

.268	.596	.319	.332	.704	.880	.444	.274	.165	.242
1.000	.349	.646	.557	.662	.395	.665	.341	.197	.189
.844	1.000	.400	.181	.363	.783	.494	.117	.230	.595
.236	.255	.156	.741	.150	.435	.665	.258	.137	.295
.274	.445	.837	.458	.388	1.000	.721	.203	.687	1.000
1.000	.759	.464	.699	.208	.201	.136	.228	.581	.486
.202	.556	.255	.207	.198	.266	.452	.172	.679	.176
.227	.781	.865	.159	.171	.707	.225	.143	.534	.300
.304	.818	.262	.174	.276	.320	.272	.821	.791	.392
.482	.214	.924	.388	.340	.965	.171	.290	.145	.168
.335	.098	.724	.139	.736	.611	.253	.323	.962	.877
.259	.884	.371	.254	.263	.611	.291	1.000	.123	.781
.998	.547	.356	.729	.200	.326	.217	.122	.379	.221
.342	.227	.236	.418	.871	.528	1.000	.349	.217	.571
.362	.270	.321	.968	.778	.234	.520	.345	.458	.539
.267	.389	.856	.167	.114	.435	.992	.152	.152	.198
.383	.384	.218	.289	.431	.335	.378	.621	.364	.310
.250	.669	.687	.261	.193	.183	.324	.951	.156	.501
.283	.330	1.000	.266	.693	.243	.446	.338	.315	.430
.565	.336	.113	.863	.124	.533	.990	.363	.301	.486
.250	.295	1.000	1.000	1.000	.242	.394	.510	.807	.508
.261	.188	.626	.272	.437	.220	.579	.195	.386	.743
.196	.528	.385	.247	.122	.280	.902	.465	.225	.504
.232	.384	.423	.761	.396	.892	.700	.575	.496	.376
.870	.244	.180	.101	.369	.277	.181	.512	.706	.749

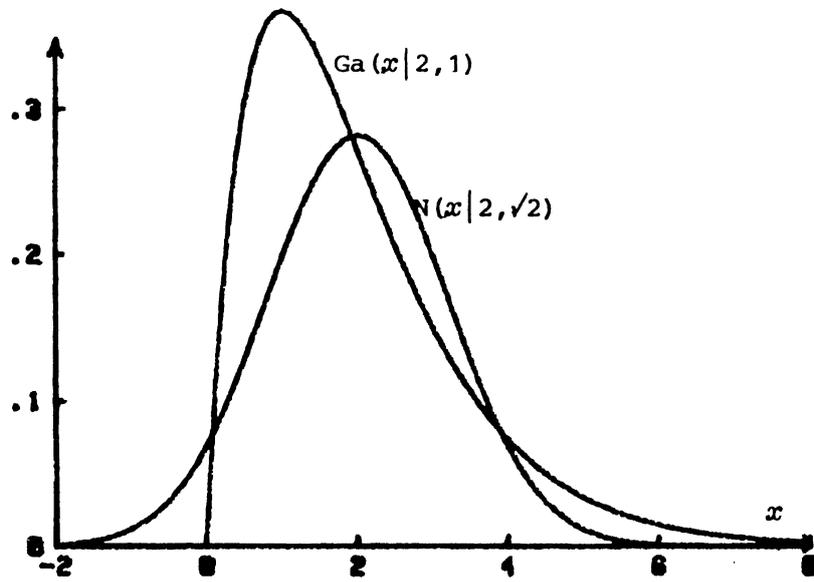


Fig. 2. Distribuciones $N(x|2, \sqrt{2})$ y $Ga(x|2, 1)$

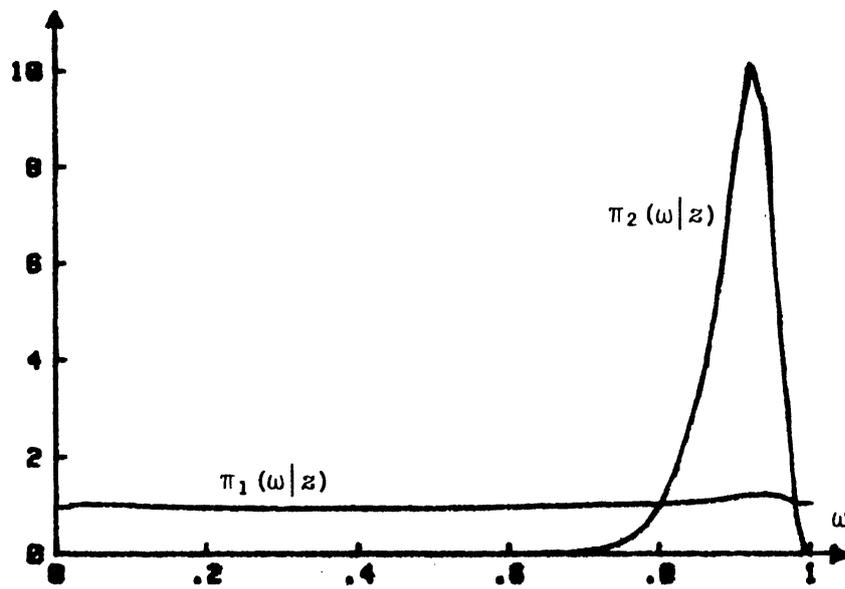


Fig. 3. Distribuciones finales del parámetro marginal

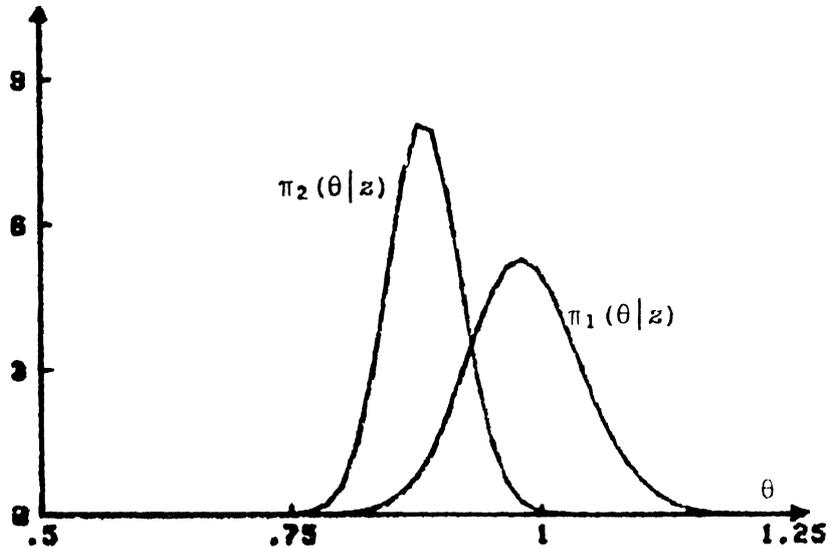


Fig. 4. Distribuciones finales del parámetro de interés

El uso de las distribuciones finales aproximadas introducidas en la sección anterior permite llegar asimismo al rechazo de la hipótesis nula para la muestra II. En efecto, los valores de $t(\omega)$ que corresponden a las modas de las distribuciones finales de ω resultan ser $t(\omega_1) = 1.047$ y $t(\omega_2) = 1.412$, que dan lugar, respectivamente, a las distribuciones finales aproximadas $\text{Ga}(\theta | 250, 262)$ y $\text{Ga}(\theta | 250, 353)$, cuyas representaciones gráficas son parecidas a las verdaderas distribuciones finales, representadas en la Figura 4.

6. DISCUSION

El método propuesto permite un análisis sistemático de posibles hipótesis sobre el modelo probabilístico que permite describir un conjunto de datos. Es importante subrayar que el método puede ser aplicado al contraste de *cualquier* distribución unidimensional. El procedimiento descrito reduce el problema original al contraste de la compatibilidad de $\theta = 1$ con la distribución final $\pi(\theta|z)$ del parámetro de interés.

La solución exacta consiste en determinar el nivel de significación necesario para incluir el punto $\theta = 1$ en un intervalo creíble o de confianza construido sobre $\pi(\theta|z)$; como ya hemos puesto de manifiesto este nivel de significación tiene una clara interpretación Bayesiana en términos de la probabilidad asignada a la hipótesis nula.

Esta solución exacta puede ser programada sin dificultades en cualquier miniordenador; no obstante, como hemos visto, es posible encontrar una excelente aproximación que permite decidir sobre la compatibilidad de los datos con la hipótesis, con la ayuda de una simple calculadora de bolsillo, y *sin necesidad de tablas especiales*.

La comparación entre modelos probabilísticos alternativos para un conjunto de datos puede llevarse a cabo mediante la comparación de las probabilidades finales a que, respectivamente, dan lugar.

El contraste de una hipótesis distribucional en la que los parámetros no han sido especificados requiere naturalmente su estimación previa a partir de los datos; en tal caso, lo que se contrasta es obviamente la compatibilidad de los datos con la distribución, totalmente especificada así obtenida. El contraste de una hipótesis distribucional con parámetros desconocidos puede llevarse a cabo sin embargo tratando a tales parámetros como parámetros marginales; este tema será tratado en un trabajo próximo.

Finalmente, sería claramente deseable disponer de un análisis comparativo, via Montecarlo, de la potencia del método de contraste sugerido con respecto a la obtenida mediante otros procedimientos presentes en la literatura, análisis que esperamos asimismo acometer en un futuro próximo.

REFERENCIAS

- BASULTO, J. (1978). Transformación raíz cuadrada versus logarítmica de una distribución gamma. *Trab. Estadist.* 29, 3, 83-88.
- BERNARDO, J.M. (1979). Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc. B* 41, 113-147 (con discusión).
- BERNARDO, J.M. (1980). A Bayesian analysis of classical hypothesis testing. *Trab. Estadist.* 31, 606-647 (con discusión).
- RENYI, A. (1976). *Cálculo de Probabilidades*. Barcelona: Reverté.