

Estadística Práctica para Astrónomos -I.
Definiciones, La Distribución Normal, Detección de
Señales. ¹

J.V. Wall

Mullard Radio Astronomy Observatory, Cavendish
Laboratory, Madingley Road,
Cambridge, CB3 0HE

¹Artículo original *Practical Statistics for Astronomers – I. Definitions, the Normal Distribution, Detection of Signal*, escrito por J.V. Wall, 1979, Quaterly Journal of the Royal Astronomical Society, Volúmen 20, página 138. Traducción al Español por José Antonio García Barreto, Instituto de Astronomía, Universidad Nacional Autónoma de México (IA-UNAM), México, D.F. Mayo 2004

Resumen

Esta es una pequeña introducción a estadística que puede ser útil para los astrónomos. Contiene definiciones, una definición de las funciones de probabilidad, una discusión de la Distribución Normal, y algunas consideraciones de detección de señales (en longitudes de onda en el óptico y en la banda de radio) con referencia en particular a ésta distribución.

1 INTRODUCCION

Los astrónomos no pueden evadir la estadística, y existen varias razones para esta situación desafortunada. La más obvia es que cada ciencia observacional es una de *probabilidades* – ninguna es más que astronomía, en la cual los astrónomos ópticos cuentan fotones individuales de objetos débiles hasta que han recibido ‘suficientes’, mientras que sus colegas radio astrónomos persisten con receptores que generan señales de ruido cuyas amplitudes son cientos de veces más grandes que las que se esperan detectar de los objetos débiles. Todos hemos recibido conocimientos de nuestros *Maestros* que ninguna cantidad determinada observacionalmente tiene ninguna utilidad a menos de que se especifique la incertidumbre asociada a la medición; esto implica que sabemos y entendemos nuestros instrumentos *y* un poco de estadística básica. También implica que otros astrónomos van a dar resultados en términos estadísticos – por ejemplo: errores estándares, límites de confiabilidad – de tal forma que en defensa propia, debemos de estar familiarizado con las implicaciones de estas estadísticas.

Después de los problemas de *detección* vienen los problemas de *muestreo* – nos encontramos frecuentemente con la necesidad de hacer deducciones acerca de varios constituyentes del Universo en base a muestras que son invariablemente pequeñas y que no es fácil hacerlas más grandes. Cómo podemos convencernos ó convencer a nuestros colegas que un efecto en nuestra muestra indica una Verdad Universal? Qué tan probable es que el efecto sea resultado de la casualidad? ó de buena/mala suerte? ó debida a la Primera Ley de Experimentación²? No siempre estamos conscientes de que una prueba apropiada

²Si el salami *puede caerse* del sandwich, lo *hará*.

existe. *Es* posible, por ejemplo, probar si el ‘grado de aglomeración’ (arbitrario y en escala no numérica) de las estructuras en una muestra de cinco fuentes de radio (ondas) está correlacionada con, digamos, latitud galáctica?

Problemas prácticos como estos (?) reciben poco tratamiento en tratados estándares de estadística (por ejemplo: Kendall & Stuart (1)). Pero la mayoría de las cosas necesariamente han sido realizadas por Aquellos Que Se Han Ido con Anterioridad, y algunos de sus resultados estan aquí reportados, en esta corta serie de artículos los cuales han sido la base de las notas preparadas para los estudiantes nuevos en investigación en el Observatorio de Radio Astronomía de Mullard (en inglés Mullard Radio Astronomy Observatory, MRAO), en el Laboratorio Cavendish, en Cambridge, Inglaterra, Cambridge, CB3 0HE.

Los artículos representan un punto de vista particular, y no me hago responsable ni de la originalidad ni de la completés. Algunos resultados estan presentados con ejemplos pero con muy poca justificación y nada de teoría la cual se puede encontrar en los libros estándares. Se incluyen tablas relacionadas al material descrito y estas se han recalculado para evitar errores de mecanografía y problemas de derechos de autor. En todos éste artículo enfatizo en dos cosas: sentido común y la necesidad de usar métodos *no-paramétricos*. No hay sustituto para el sentido común. La segunda es a fuerzas: las pruebas paramétricas usuales suponen un conocimiento de las distribuciones probabilísticas de las variables que estamos muestreando. (ejemplos son la prueba t de Estudiantes, la prueba F, y el coeficiente de correlación estándar r , las cuales presuponen la Distribución Normal para las variables involucradas.) Raramente tenemos ese lujo, y de hecho, la distribución de probabilidad involucrada es frecuentemente el objetivo de nuestra investigación. Por lo tanto debemos utilizar pruebas *no-paramétricas*. Los libros usualmente no hacen la distinción, y si lo hacen, el tratamiento no paramétrico es muy común que sea sólo menos de uno por ciento del volúmen. EL monograma de Siegel (2) es una excepción admirable y se hacen referencia a él después.

En éstos artículos discutiré cuatro tópicos que muy probablemente representan los más comunmente utilizados con estadísticas por los astrónomos:

1. Detección de señales: ¿Cómo se realizó? ¿Puede uno creela?

2. Correlación: ¿Existe? ¿A qué nivel de confianza ó significado?
3. Estimación Paramétrica de un Modelo: ¿cuáles son los errores apropiados de los parámetros? ¿Ha sido el modelo y/ó la hipótesis razonable en primera instancia?
4. Comparación de muestras: (a) con las predicciones de un modelo y/ó hipótesis, y (b) una con otra; ¿pertenecen a la misma población (ó conjunto universal)? ¿Difieren significativamente? ¿Difieren en un sentido predicho?

Como un preámbulo inevitable tengo que presentar algunas definiciones y distribuciones (Sección 2) las cuáles ciertamente se pueden encontrar en la literatura, ya sea astronómica y estadística, y discutiré la distribución Normal (Sección 3). El tópico (1) está considerado en la Sección 4, y los demás tópicos en los artículos subsecuentes.

2 DEFINICIONES

Estadístico(a): es ‘una persona preparada para estimar la probabilidad de que salga el Sol basada en movimientos anteriores’(3). Esta definición, que admito que he tomado fuera de contexto, fué dada por un eminente estadístico; puede confirmar nuestra peor suspicacia acerca de eminentes estadísticos ó algo parecido. Por lo menos, debe enfatizar el ser cuidadoso con las definiciones – ¿que es mañana si no sale el Sol? Seamos más cuidadosos:

Estadística es un término vagamente utilizado para describir tanto a la ciencia como los valores. De hecho la ciencia es

Inferencia Estadística: la determinación de propiedades de la *población ó conjunto universal* a partir de la *muestra*,

mientras que

Estadística son los valores (usualmente, pero no necesariamente, numéricos) determinados para algunos ó todos los valores de la muestra.

Buenas Estadísticas son aquéllas de la cuales nuestras conclusiones concernientes a la *población ó conjunto universal* son estables de una muestra con otra muestra, mientras

que *buenas muestras* proveen buenas estadísticas y requieren un diseño apropiado del experimento.

Lo *bueno* del experimento, la muestra o la estadística está indicado por el

Nivel de Significado: supongamos que realizamos un experimento para distinguir entre dos hipótesis que rivalizan una con otra, la *hipótesis nula* (H_0 ; ‘falla’, no resultado, no detección, no correlación) y su alternativa (H_1 ; ‘éxito’, etc.). Antes del experimento nos aseguramos de estar familiarizados con la ‘falla’ determinando, *suponiendo que H_0 es verdadera*, el conjunto de todos los valores de la estadística en la prueba, es decir, la estadística que hemos determinado utilizar para decidir entre H_0 y H_1 . Además, suponemos que cuando hacemos el experimento obtenemos un valor de la estadística que es *inusual* en comparación con este conjunto, tan inusual que, digamos, sólo uno por ciento de todos los valores calculados bajo la hipótesis H_0 sean tan extremos. Podemos entonces rechazar H_0 en favor de H_1 al uno por ciento de *nivel de confiabilidad ó significado*. El *nivel de confiabilidad ó significado* es por lo tanto la probabilidad de rechazar H_0 cuando, es de hecho, cierta.

Ahora consideremos N valores de x_i donde $i = 1, 2 \dots N$ y x pueden tener una distribución discreta ó continua. Las siguientes definiciones son en general:

1. Mediciones de localización

Promedio Aritmético³:
$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

Valor Medio: Dado un conjunto de números x_i , el valor medio se calcula de la siguiente manera. Primero se hace una lista de los números x_i de tal forma que van del menor valor al máximo valor y se les abjudica el nuevo orden (digamos x_j , donde j va desde uno hasta el valor final N igual que los valores iniciales de x_i). El valor medio puede tomar dos valores dependiendo de si el número N es entero par ó impar:

³Nota del traductor: existen varias notaciones para indicar el promedio de una cantidad, por ejemplo si la cantidad son las posiciones en el eje x de varios objetos, que denotamos como x , su promedio puede estar indicado como: (1) $\langle x \rangle$, (2) μ , (3) \bar{x} , (4) \hat{x} etcétera. La notación $\sum_{i=1}^N x_i$ es equivalente a realizar la suma de N valores de la posición x iniciando desde x_1 hasta x_N , es decir: $x_1 + x_2 + x_3 + x_4 + \dots + x_N$

$\mathbf{x}_{\text{medio}} = \mathbf{x}_j$ donde $j = N/2 + 0.5$ en el caso de que N sea impar⁴.

$\mathbf{x}_{\text{medio}} = \frac{1}{2}(\mathbf{x}_j + \mathbf{x}_{j+1})$ donde $j = N/2$ en el caso de que N sea par⁵.

Valor más frecuente: x_{frec} es el valor de x_i que ocurre con más frecuencia.

2. Medidas de Dispersión (ó valores que difieren del promedio aritmético)

Desviación Media: $\overline{\Delta x} = \frac{1}{N} \sum_{i=1}^N |x_i - x_{\text{medio}}|$

Varianza ó Desviación Media Cuadrada: $\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$

Desviación Estándar: $\sigma = \sqrt{\sigma^2}$

3. Momentos

n-ésimo: $u_n = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^n$

(Los momentos pueden ser tomados para *cualquier* valor de x ; aquéllos alrededor del promedio aritmético como en el párrafo anterior se denominan *momentos centrales*. Nótese que $\mu_2 = \sigma^2$; éste y los siguientes momentos caracterizan una función de probabilidad (definida en el siguiente párrafo). Los primeros dos momentos no tienen uso, ya que $\mu_0 \equiv 1$ y $\mu_1 \equiv 0$.

Distorción: $\beta_1 = \mu_3^2 / \mu_2^3$ indica desviación de la simetría;

$\beta_1 = 0$ para simetría alrededor del promedio aritmético, μ .

Concentración: $\beta_2 = \mu_4 / \mu_2^2$ indica grado de concentración alrededor del promedio aritmético, μ , es decir, indica si hay un pico,

$\beta_2 = 3$ para una distribución Normal.

⁴Nota del traductor: ejemplo: si se tienen 5 valores x_1, x_2, x_3, x_4 y x_5 , el valor medio, x_{medio} , será el valor dado por x_3

⁵Nota del traductor: ejemplo: si se tienen 8 valores $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ y x_8 , el valor medio, x_{medio} , será el valor dado por $\frac{1}{2}(x_4 + x_5)$

Finalmente considérese *distribuciones de probabilidad*: si x es una variable azarosa y continua, entonces $f(x)$ es su *densidad de función de probabilidad* si cumple con los siguientes requisitos:

- (1) Probabilidad ($a \leq x \leq b$) = $\int_a^b f(x) dx$
- (2) $\int_{-\infty}^{\infty} f(x) dx = 1$
- y (3) $f(x)$ es un valor único, no negativo, dado cualquier valor de x .

La *función de distribución* correspondiente es: $F(x) = \int_{-\infty}^x f(y) dy$.

Densidades de probabilidad y funciones de distribución pueden ser similarmente definidas para conjuntos discretos de valores $x = x_1, x_2, x_3, x_4, \dots, x_n$, y para funciones multivariable. Las funciones más conocidas aparecen en la Tabla 1, junto con sus medidas de localización y dispersión – nótese que las definiciones anteriores para estas funciones pueden ser escritas en una forma de integral (matemática) para funciones continuas. La tabla incluye algunas sugerencias de cómo y dónde se origina cada distribución, y para la mayoría de ellas evito una mayor explicación. Pero existe una cuyo papel es fundamental y que no puede ser tratada de esa manera tan superflua. Esta es, por supuesto:

3 LA DISTRIBUCION GAUSSIANA NORMAL

para la cual la función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2}.$$

Figure 1 muestra la apariencia bien conocida de esta función y la de su integral, la función de distribución correspondiente. Ambas funciones estan tabuladas in Tabla II. Dados los símbolos que se escogieron, el promedio es μ , la varianza es σ^2 y la dispersión estándar es σ .

TABLE I
The common probability distributions

Distribution	Density function	Mean	Variance	Raison d'Être
Uniform	$f(x; a, b) = \frac{1}{b-a}$ $a \leq x \leq b$ $= 0$ $x < a, > b$	$(a+b)/2$	$(b-a)^2/12$	In the study of rounding errors, and as a tool in theoretical studies of other continuous distributions.
Binomial	$f(x; n, p, q) = \binom{n}{x} p^x q^{n-x}$	np	npq	x is the number of 'successes' in an experiment with two possible outcomes, one ('success') of probability p , and the other ('failure') of probability $q = 1 - p$. Becomes a Normal distribution as $n \rightarrow \infty$.
Poisson	$f(x; \mu) = e^{-\mu} \mu^x / x!$	μ	μ	The limit for the Binomial distribution as $p \ll 1$, setting $\mu = np$. It is the 'count-rate' distribution, e.g. take a star from which an average of μ photons are received per Δt (out of a total of n emitted, $p \ll 1$); the probability of receiving x photons in a Δt is $f(x; \mu)$. Tends to the Normal distribution as $\mu \rightarrow \infty$.
Normal (Gaussian)	$f(x; \mu, \sigma) = [1/(2\pi\sigma^2)] \exp[-(x-\mu)^2/2\sigma^2]$	μ	σ^2	The essential distribution; see text. Central Limit Theorem ensures that majority of 'scattered things' are dispersed according to $f(x; \mu, \sigma)$.
Chi-square	$f(x^2; \nu) = \frac{x^{\nu/2-1} \exp(-x^2/2)}{2^{\nu/2} \Gamma(\nu/2)}$	ν	2ν	Vital in the comparison of samples, model testing; characterizes the dispersion of observed samples from the expected dispersion, because if x_i is a sample of ν variables Normally and independently distributed with means μ_i and variances σ_i^2 , then $\chi^2 = \sum_{i=1}^{\nu} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$ obeys $f(x^2; \nu)$.
Student t	$f(t; \nu) = \Gamma((\nu+1)/2) \cdot (1+t^2/\nu)^{-(\nu+1)/2} / (\pi\nu) \Gamma(\nu/2)$	0	$\nu(\nu-2)$ (for $\nu > 2$)	Invariably tabulated and used in integral form. Tends to Normal distribution as $\nu \rightarrow \infty$.
F	$f(F; \nu_1, \nu_2) = \frac{\Gamma((\nu_1 + \nu_2)/2)}{\Gamma(\nu_1/2) \Gamma(\nu_2/2)} \cdot (\nu_1/\nu_2)^{\nu_1/2} \cdot \frac{F^{\nu_1/2-1}}{(1 + F\nu_1/\nu_2)^{\nu_1/2} (1 + F)^{\nu_2/2}}$	$\frac{\nu_1(\nu_2-2)}{\nu_1(\nu_2-2) + \nu_2}$ (for $\nu_2 > 2$)	$\frac{2\nu_1\nu_2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_1 - 4)}$ (for $\nu_1 > 4$)	For comparison of means, Normally-distributed populations; if n x_{1j} are taken from a Normal population (μ_1, σ_1), and if n x_{2j} are found as in text, then $t = \sqrt{n}(\bar{x}_1 - \mu_1)/\sigma_1$ is distributed as $f(t; \nu)$ where 'degrees of freedom' $\nu = n - 1$. Student's t can also be formulated to compare means for samples from Normal populations with same σ , different μ (4). Tends to Normal as $\nu \rightarrow \infty$.
				For comparison of two variances, or of more than two means; if two statistics (χ_1^2 and χ_2^2) each follow the Chi-square distribution, then $F = \frac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2}$ is distributed as $f(F; \nu_1, \nu_2)$. Care required in application; see (4), (9).

Figure 1: Tabla I.

TABLE II
The normal distribution

$$t = \frac{x - \mu}{\sigma} \quad f(t) = \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\} \quad F(t) = \int_{-\infty}^t f(y) dy$$

t	f(t)	F(t)	t	f(t)	F(t)	t	f(t)	F(t)	t	f(t)	F(t)
0.00	0.3989	0.5000	0.00	0.3989	0.5000	1.00	0.2420	0.8413	1.00	0.2420	0.8413
0.01	0.3989	0.5040	0.01	0.3989	0.5040	1.01	0.2420	0.8438	1.01	0.2420	0.8438
0.02	0.3989	0.5080	0.02	0.3989	0.5080	1.02	0.2420	0.8463	1.02	0.2420	0.8463
0.03	0.3989	0.5120	0.03	0.3989	0.5120	1.03	0.2420	0.8488	1.03	0.2420	0.8488
0.04	0.3989	0.5160	0.04	0.3989	0.5160	1.04	0.2420	0.8513	1.04	0.2420	0.8513
0.05	0.3989	0.5200	0.05	0.3989	0.5200	1.05	0.2420	0.8538	1.05	0.2420	0.8538
0.06	0.3989	0.5240	0.06	0.3989	0.5240	1.06	0.2420	0.8563	1.06	0.2420	0.8563
0.07	0.3989	0.5280	0.07	0.3989	0.5280	1.07	0.2420	0.8588	1.07	0.2420	0.8588
0.08	0.3989	0.5320	0.08	0.3989	0.5320	1.08	0.2420	0.8613	1.08	0.2420	0.8613
0.09	0.3989	0.5360	0.09	0.3989	0.5360	1.09	0.2420	0.8638	1.09	0.2420	0.8638
0.10	0.3989	0.5400	0.10	0.3989	0.5400	1.10	0.2420	0.8663	1.10	0.2420	0.8663
0.11	0.3989	0.5440	0.11	0.3989	0.5440	1.11	0.2420	0.8688	1.11	0.2420	0.8688
0.12	0.3989	0.5480	0.12	0.3989	0.5480	1.12	0.2420	0.8713	1.12	0.2420	0.8713
0.13	0.3989	0.5520	0.13	0.3989	0.5520	1.13	0.2420	0.8738	1.13	0.2420	0.8738
0.14	0.3989	0.5560	0.14	0.3989	0.5560	1.14	0.2420	0.8763	1.14	0.2420	0.8763
0.15	0.3989	0.5600	0.15	0.3989	0.5600	1.15	0.2420	0.8788	1.15	0.2420	0.8788
0.16	0.3989	0.5640	0.16	0.3989	0.5640	1.16	0.2420	0.8813	1.16	0.2420	0.8813
0.17	0.3989	0.5680	0.17	0.3989	0.5680	1.17	0.2420	0.8838	1.17	0.2420	0.8838
0.18	0.3989	0.5720	0.18	0.3989	0.5720	1.18	0.2420	0.8863	1.18	0.2420	0.8863
0.19	0.3989	0.5760	0.19	0.3989	0.5760	1.19	0.2420	0.8888	1.19	0.2420	0.8888
0.20	0.3989	0.5800	0.20	0.3989	0.5800	1.20	0.2420	0.8913	1.20	0.2420	0.8913
0.21	0.3989	0.5840	0.21	0.3989	0.5840	1.21	0.2420	0.8938	1.21	0.2420	0.8938
0.22	0.3989	0.5880	0.22	0.3989	0.5880	1.22	0.2420	0.8963	1.22	0.2420	0.8963
0.23	0.3989	0.5920	0.23	0.3989	0.5920	1.23	0.2420	0.8988	1.23	0.2420	0.8988
0.24	0.3989	0.5960	0.24	0.3989	0.5960	1.24	0.2420	0.9013	1.24	0.2420	0.9013
0.25	0.3989	0.6000	0.25	0.3989	0.6000	1.25	0.2420	0.9038	1.25	0.2420	0.9038
0.26	0.3989	0.6040	0.26	0.3989	0.6040	1.26	0.2420	0.9063	1.26	0.2420	0.9063
0.27	0.3989	0.6080	0.27	0.3989	0.6080	1.27	0.2420	0.9088	1.27	0.2420	0.9088
0.28	0.3989	0.6120	0.28	0.3989	0.6120	1.28	0.2420	0.9113	1.28	0.2420	0.9113
0.29	0.3989	0.6160	0.29	0.3989	0.6160	1.29	0.2420	0.9138	1.29	0.2420	0.9138
0.30	0.3989	0.6200	0.30	0.3989	0.6200	1.30	0.2420	0.9163	1.30	0.2420	0.9163
0.31	0.3989	0.6240	0.31	0.3989	0.6240	1.31	0.2420	0.9188	1.31	0.2420	0.9188
0.32	0.3989	0.6280	0.32	0.3989	0.6280	1.32	0.2420	0.9213	1.32	0.2420	0.9213
0.33	0.3989	0.6320	0.33	0.3989	0.6320	1.33	0.2420	0.9238	1.33	0.2420	0.9238
0.34	0.3989	0.6360	0.34	0.3989	0.6360	1.34	0.2420	0.9263	1.34	0.2420	0.9263
0.35	0.3989	0.6400	0.35	0.3989	0.6400	1.35	0.2420	0.9288	1.35	0.2420	0.9288
0.36	0.3989	0.6440	0.36	0.3989	0.6440	1.36	0.2420	0.9313	1.36	0.2420	0.9313
0.37	0.3989	0.6480	0.37	0.3989	0.6480	1.37	0.2420	0.9338	1.37	0.2420	0.9338
0.38	0.3989	0.6520	0.38	0.3989	0.6520	1.38	0.2420	0.9363	1.38	0.2420	0.9363
0.39	0.3989	0.6560	0.39	0.3989	0.6560	1.39	0.2420	0.9388	1.39	0.2420	0.9388
0.40	0.3989	0.6600	0.40	0.3989	0.6600	1.40	0.2420	0.9413	1.40	0.2420	0.9413
0.41	0.3989	0.6640	0.41	0.3989	0.6640	1.41	0.2420	0.9438	1.41	0.2420	0.9438
0.42	0.3989	0.6680	0.42	0.3989	0.6680	1.42	0.2420	0.9463	1.42	0.2420	0.9463
0.43	0.3989	0.6720	0.43	0.3989	0.6720	1.43	0.2420	0.9488	1.43	0.2420	0.9488
0.44	0.3989	0.6760	0.44	0.3989	0.6760	1.44	0.2420	0.9513	1.44	0.2420	0.9513
0.45	0.3989	0.6800	0.45	0.3989	0.6800	1.45	0.2420	0.9538	1.45	0.2420	0.9538
0.46	0.3989	0.6840	0.46	0.3989	0.6840	1.46	0.2420	0.9563	1.46	0.2420	0.9563
0.47	0.3989	0.6880	0.47	0.3989	0.6880	1.47	0.2420	0.9588	1.47	0.2420	0.9588
0.48	0.3989	0.6920	0.48	0.3989	0.6920	1.48	0.2420	0.9613	1.48	0.2420	0.9613
0.49	0.3989	0.6960	0.49	0.3989	0.6960	1.49	0.2420	0.9638	1.49	0.2420	0.9638
0.50	0.3989	0.7000	0.50	0.3989	0.7000	1.50	0.2420	0.9663	1.50	0.2420	0.9663

Figure 0: Tabla II.

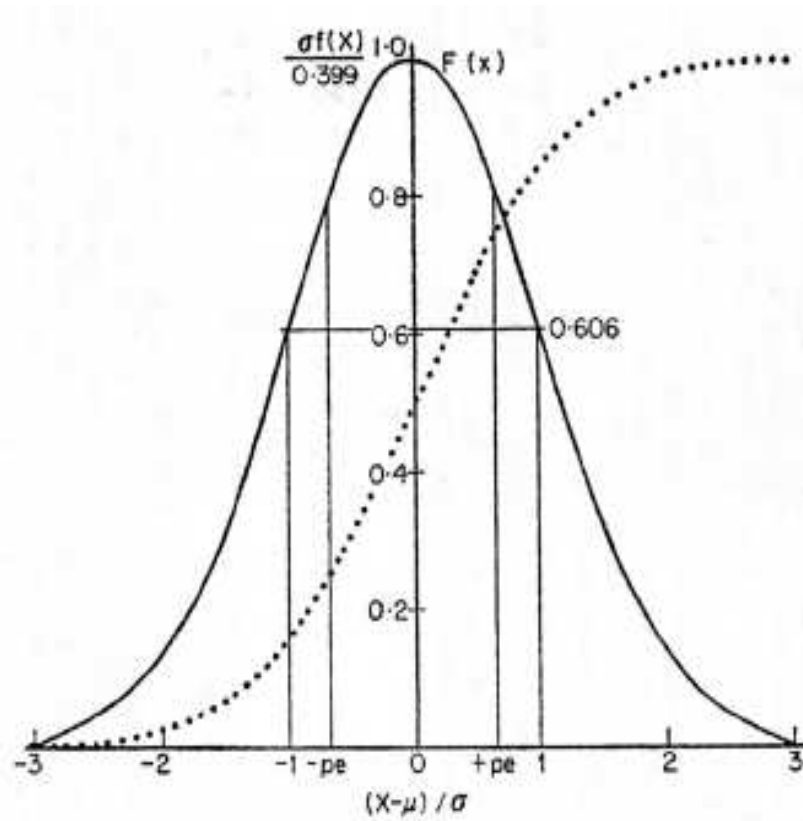


Figure 1: La Distribución Normal. La función de probabilidad de densidad (curva continua) es $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2}$ y la función de distribución (curva puntuada) es $F(x) = \int_{-\infty}^x f(y) dy$. 50 por ciento del área cae entre $\pm ep$ (error probable). Ver Tablas II y III.

Digamos que escogemos una muestra de N valores de x , sean x_i con i desde 1 hasta N , de una población de x de las cuales creemos que están distribuidas de acuerdo a una distribución Normal – el argumento de tal creencia puede esperar un párrafo. ¿Cómo es que estimamos el promedio aritmético de la población, μ , y su varianza, σ^2 ? No lo hacemos necesariamente aplicando las definiciones de la Sección 2 a nuestros valores de x_i porque estos pueden no resultar en estimaciones *sin prejuicio*. De hecho, se puede mostrar (véase referencia (4)) que para una población distribuida en una forma Normal, el promedio aritmético de la muestra, \bar{x}_m , es en verdad una estimación *sin prejuicio* de μ , pero la estimación *sin prejuicio* de la varianza σ^2 es (la varianza de la muestra que está dada por):

$$\sigma_m^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_m)^2$$

la cual **difiere** de la definición de σ^2 dada anteriormente por el factor $(N/(N-1))$.

Este factor puede dar lugar a confusiones: a σ_m^2 se le conoce como la **varianza de la muestra**, mientras que σ^2 definida en la Sección 2 se le denomina la varianza de la población (conjunto universo) y, por supuesto, $\sigma_m^2 \rightarrow \sigma^2$ cuando N tiende a ser muy muy grande ($N \rightarrow \infty$). (El que debe haber una diferencia es fácil de entender: los valores de x_i de nuestra muestra son utilizados para encontrar \bar{x}_m , es decir, una estimación del promedio aritmético μ , y aunque este es un valor *sin prejuicio* estimado de μ , es la estimación la que da un valor *mínimo* de la suma de los cuadrados de las desviaciones de la muestra, y por lo tanto nos da una estimación *baja* de la varianza. La teoría provee el factor apropiado de corrección, léase $N/(N-1)$.)

La desviación estándar (error) en \bar{x}_m , nuestra estimación de μ , es σ_m/\sqrt{N} . Aún más, si tenemos n estimaciones de μ , digamos \bar{x}_j (donde j vá desde uno hasta n) cada una teniendo un error asociado σ_j , la **mejor** estimación de μ es el *promedio sopesado*, digamos :

$$\bar{x}_w = \frac{\sum_{j=1}^n w_j x_j}{\sum_{j=1}^n w_j}$$

donde los pesos asignados a cada valor, w_j , están dados por:

$$w_j = \frac{1}{\sigma_j^2},$$

es decir, son los valores recíprocos de las varianzas de la muestra. La mejor estimación de la varianza de \bar{x}_w es:

$$\sigma_w^2 = \frac{1}{n \sum_{j=1}^n 1/\sigma_j^2}$$

Pero regresando a la distribución Normal por sí misma; ¿porqué se le considera Mágica? ¿Porqué esperamos poblaciones (universo), conjuntos de datos, mediciones, etc. que la cumplan? Sabemos por experiencia que sí la cumplen con frecuencia. Una obvia razón para la prominencia de la distribución Normal es que dos distribuciones que se encuentran frecuentemente en la Naturaleza, son la Binomial y la de Poisson (ver Tabla 1), cada una tiende a la Normal en sus límites respectivos cuando el número de valores, n es muy muy grande y tiende a infinito, $n \rightarrow \infty$ y $\mu \rightarrow \infty$. Pero la verdadera razón quizás se base en que quizás el teorema más importante en Inferencia Estadística, Teoría de Medición y Diseño Experimental (veáse referencias [1] y [4]),

El Teorema del Límite Central: supóngase que variables independientes azarosas X_i con función de densidad de probabilidad DESCONOCIDA están idénticamente distribuidas con un valor promedio aritmético μ y varianza σ^2 (ambos valores finitos). Tómesese una muestra de n de estas variables; mientras más grande sea el valor de n , la distribución del valor promedio aritmético de X , \bar{X} , de la muestra tiende a una distribución Normal con promedio μ , mientras que la varianza de la muestra (σ_m^2) tiende a σ^2/n .

La tendencia de las distribuciones Binomial y de Poisson a la Normal es simplemente un ejemplo particular de la operación de este teorema milagroso. En verdad, el teorema aplica a variables distribuidas tanto en forma discreta como en forma continua y bajo ciertas condiciones débiles es aún posible relajar los requisitos de independencia y de distribuciones idénticas para las X_i ; también se encuentran distribuciones Normales para los promedios (y aún para otras combinaciones lineales de las variables).

El Teorema del Límite Central permite estimar probabilidades cuantitativas en situaciones experimentales donde la forma de la distribución de probabilidad es desconocida. En particular, los miembros individuales de muchos conjuntos de datos que adquirimos son el resultado de alguna clase de promedio – ¡aún si lo hacemos "a ojo"!– y en tales circunstancias el Teorema del Límite Central nos dice que los datos tienden a una distribución Normal. Como resultado, podemos (y lo hacemos) estimar los parámetros de la población (universo) μ y σ^2 a partir de muestras sencillas ó a partir de la combinación de resultados de diferentes muestras, en la manera descrita anteriormente, con la certeza en el pensamiento de que el Teorema del Límite Central tiene (probablemente) justificado su procedimiento.

Por ejemplo, deseamos determinar una intensidad a partir de varias mediciones: la mejor estimación es por lo tanto $\bar{x}_m \pm \sigma_m$ tal y como se definió anteriormente.

Alternativamente, dadas varias estimaciones de la intensidad, digamos $\bar{x}_j \pm \sigma_j$, la mejor estimación⁶ es $\bar{x}_w \pm \sigma_w$ tal como se muestra en los párrafos anteriores.

Pero el Teorema del Límite Central no siempre nos dá una distribución Normal, como veremos al discutir la detección de señales. Por supuesto, existe un punto de precaución antes de que llegemos alla. Debido a la prominencia de la distribución Normal, nos dá una manera útil de describir errores, dispersiones (ó desviaciones del promedio), ó diferencias, las cuales son intuitivamente muy aceptables. Utilizamos la distribución Normal para este propósito al mencionar errores ó diferencias en términos de la desviación estándar, la cual para una distribución Normal es la raíz cuadrada del cuadrado de la desviación, σ_m (ó diferencia con respecto al promedio aritmético) "corregida" por el factor $(\sqrt{N/[N-1]})$; para la distribución Normal la región $(\bar{x} - \sigma) \leq x \leq (\bar{x} + \sigma)$ contiene 68.3 por ciento del área bajo la curva, es decir, 68.3% de los "valores esperados" El punto de precaución es que *la distribución Normal tiene alas muy cortas* – es decir, sus valores decaén muy rápidamente más allá de $\pm 2\sigma_m$, como se muestra en la Figura 1. Considérese, por ejemplo,

⁶Si se conoce *a priori* que la frecuencia de distribución de intensidades no es uniforme (Tabla 1), entonces se puede argumentar que \bar{x}_w no se determina *sin prejuicio* aún y cuando los errores obedecen una distribución Normal precisamente. Se necesitan hacer cambios apropiados (ver referencia [5]). Este es un ángulo de un argumento de discusión entre dos tipos de estadísticos, los Bayesianos y los Otros, que lleva facilmente 200 años. Dejemosla en paz, al menos por el momento.

Table 3: Tabla III. *Las alas de la Distribución Normal*. Porcentaje del área bajo la curva de la Distribución Normal en las diferentes regiones

m	$\geq m\sigma$ (una ala)	$\geq m\sigma, \leq -m\sigma$ (ambas alas)	$-m\sigma \leq x \leq m\sigma$ (entre las alas)
0.0	50.0000000	100.0000000	0.0000000
0.5	30.8500000	61.7100000	38.2900000
1.0	15.8700000	31.7300000	68.2700000
1.5	6.6810000	13.3600000	86.6400000
2.0	2.2750000	4.5500000	95.4500000
2.5	0.6210000	1.2400000	98.7600000
3.0	0.1350000	0.2700000	99.7300000
3.5	0.0233000	0.0465000	99.9540000
4.0	0.0031700	0.0063300	99.9937000
4.5	0.0003400	0.0006800	99.9993200
5.0	0.0000287	0.0000573	99.9999430

la comparación de dos conjuntos independientes de mediciones (ficticias de la posición en el cielo de objetos celestes) de ascensión recta como se muestra en la Figura 2. Si las diferencias estuvieran distribuidas en una manera Normal, un procedimiento válido sería calcular su *promedio aritmético* para determinar la diferencia sistemática, y después calcular la varianza de la muestra, $\sigma_{m-diff}^2 = (N/[N - 1])$ veces la desviación media cuadrada. Supóngase que σ_1^2 y σ_2^2 son las varianzas conocidas para cada una de los conjuntos independientes de ascensión recta; la varianza que se espera obtener en la diferencia (ver por ejemplo la referencia [9]) es $\sigma_1^2 + \sigma_2^2$ y la varianza de la diferencia calculada, σ_{m-diff}^2 puede compararse con ésta, quizás para verificar las estimaciones de las varianzas individuales. Sin embargo, la Fig.2 muestra algunas grandes diferencias del tipo que invariablemente parece ocurrir en comparaciones de este tipo y que extiende las *alas* más allá de aquellas esperadas para una distribución Normal. (Tales diferencias pueden originarse de errores computacionales ó debido a algunas variaciones de la distribución de probabilidad considerada, ver Sección 4.) Aún más, vale la pena hacer notar lo que sucede si calculamos nuestras estimaciones del promedio y varianza de acuerdo a las distribuciones Normales – éstas cantidades estan completamente dominadas por los valores no esperados ó anormales, y la distribución Normal que ellos implican (la curva punteada) dá un ajuste (ó modelo) a los datos muy malos. Pero todavía podemos aún describir éstos en términos de una distribución Normal si somos lo suficientemente cuidadosos. Una manera es el descartar los valores anormales (valores grandes que hacen muy grandes las alas de la curva) antes de calcular nuestras estimaciones de nuestra media y varianza; pero quizás esto no sea posible y es **sin duda muy peligroso**. Un método más seguro es utilizar la *mediana* como una medida de localización, y adoptar como medida de dispersión los valores de $\pm\Lambda$ (depués de substraer el desplazamiento relativo) dentro de los cuales los valores caén 68.3 por ciento. Esta última estadística es muy comunmente utilizada y, por supuesto, en analogía con la distribución Normal, corresponde exactamente a $\pm 1\sigma$.

En este respecto, *es muy importante estar consientes del error probable*. Este es $\pm\Lambda$ dentro del cual están 50 por ciento de los valores (y no el 68.3 por ciento). Es algunas veces utilizado con justificación cuando los errores en un experimento no están distribuidos de una manera Normal. Sin embargo, es menor que σ ($\sim 0.7\sigma$) y por lo tanto su utilización

proveé un mejoramiento cosmético a los resultados. Esta última tentación algunas veces es muy atractiva para algunos experimentalistas.

4 DETECCION DE SEÑALES

La Fig. 3(a) muestra datos obtenidos al "juntar" 50 rastreos de una señal de la emisión (conocida como emisión de continuo en ondas de radio) en declinación de un objeto celeste con el radio telescopio de 64m en Parkes, Australia operando en el rango de las microondas a una frecuencia de 5 GHz. Representa un intento de detección de la emisión de un par de galaxias conocidas como NGC 3309 y NGC 3311 cuyas posiciones en declinación están indicadas. Un aspecto importante e indiscutible de la observación es que existe mucho ruido. Pero ¿hay señal? La búsqueda de líneas de emisión de moléculas en el rango de las microondas es un problema análogo; la gráfica de la salida de muchos canales de un correlador digital puede ser muy similar a la Fig. 3(a) pero en vez de tener desplazamiento en distancia en el eje horizontal, se tendría desplazamiento en frecuencia. ¿existe monóxido de carbono a la frecuencia esperada? Espectrofotometría digital provee datos muy similares y problemas de detección muy similares en astronomía óptica.

Éxito en tales experimentos se describe usualmente como el obtener una detección a un nivel de 2σ , a un nivel de 3σ , etc. Esto es, por supuesto, haciendo una alusión directa a la distribución Normal; éxito a un nivel de 2σ significa que existe una señal presente con una amplitud de dos veces la desviación estándar del "ruido". Pero ¿porqué podemos hacer referencia a la distribución Normal para describir la gráfica de nuestro experimento con las altas y bajas de amplitud? Es por el simple hecho de que *podemos* razonablemente esperar que estas fluctuaciones se comporten aproximadamente como una distribución Normal. En radio astronomía el Teorema del Límite Central hace el trabajo: las señales débiles detectadas por un radio telescopio tiene fluctuaciones dominantes sobrepuestas que se originan en el ruido de voltajes en las primeras etapas de los receptores de señales. Pero cada dato que sale del receptor es el resultado de algun promedio *a través de* nuestra "salida de tiempo constante" (un filtro que deja pasar sólo señales de baja frecuencia), el cual produce un conjunto de fluctuaciones muy cercanas a una distribución Normal (dado

nuestro conjunto *estable* de receptores y atmósfera). (En nuestras observaciones de la Fig. 3(a), la señal (original) de salida del receptor estuvo rápidamente muestreada en intervalos de 10 milisegundos y cada punto mostrado es el promedio de muchos de esos muestreos.) En astronomía óptica, el límite de la distribución de Poisson hace el trabajo: astrónomos ópticos sufren de la falta de fotones, pero aquellos que sí reciben son muy intensos, de tal forma que las fluctuaciones de la salida no reflejan las características del receptor sino más bien la estadística de Poisson de los "eventos ó fotones" que arriban. Tan pronto como ha existido suficiente tiempo de integración para recibir más que el promedio de 10 eventos por intervalo de muestreo, estamos hablando de fluctuaciones de muestreos distribuidos aproximadamente Normales. (Esto puede verse como un ejemplo particular de la operación del Teorema del Límite Central, que difiere del caso del radio astrónomo solamente en el hecho de que la distribución de probabilidad es *conocida*.)

En este aspecto, entonces, consideremos el más común de los éxitos del astrónomo el famoso resultado de 2σ . La Tabla III da el área bajo las alas de la curva Normal más allá de varios múltiplos de σ , y nos dice que 4.6 por ciento del área (total) está fuera del límite de $\pm 2\sigma$. Un resultado *de* 2σ , por lo tanto, ocurre *de pura chiripa ó en una forma azarosa* 4.6 veces por cada 100. En la mayoría de los casos sabemos el signo del resultado (si es que hay tal resultado) – las transiciones atómicas ó moleculares son predichas en emisión ó en absorción (signo positivo ó negativo respectivamente); los cuerpos celestes emiten solamente flujo (de radiación) positiva. Así las cosas, la probabilidad de un evento azaroso se convierte en el área bajo la curva de una de las alas y baja a 2.3 eventos de cada 100 si ocurrieran en forma azarosa ó dicho de otra manera sería un evento de cada 50. El resultado genuino de 2σ es por lo tanto creíble, y usualmente uno ó dos factores pueden ser dignos de análisis. ¿ Es el ancho de la línea el ancho esperado, es decir, el ancho de la línea atómica, ó molecular, predicho, ó representa el ancho de la respuesta del radio telescopio? ¿ Qué se puede decir de la coincidencia con respecto a la posición ó frecuencia de la onda predicha? Un resultado genuino de 2σ que satisface una ó dos más de tales consideraciones es digna de ser tomada en cuenta en forma seria.

Pero no lo hacemos, y de hecho, astrónomos sienten la necesidad de dar resultados de 3σ ó hasta 7σ , para las cuales las probabilidades de eventos azarosos son exageradamente

pequeños (por ejemplo 1.3 eventos de cada 1000 para un resultado de 3σ). El problema se origina por varias razones: (a) *todos* los observadores están prejuiciados (de otra forma nunca nos molestaríamos en observar), y b) *nadie* puede estimar la dispersión estándar 1σ del ruido en una forma apropiada. La primera razón (a) en su máxima expresión insinúa –¿cuántas veces realizó el experimento el astrónomo antes de obtener el resultado particular que reporta? O en forma alternativa, ¿hizo sólo una observación terminando la integración cuando se obtuvo el resultado "esperado"? (Nótese que este último punto constituye un argumento razonable en contra del uso de los despliegues de datos en-línea ó interactivos en la pantalla de la computadora.) Ambas frases llevan, por supuesto, a un significado sobreestimado. La razón (b) puede ser considerada con más empatía. Durante cualquier observación algo (ejemplo nubes, interferencia) podría entorpecer la distribución de probabilidad por un período de tiempo, aunque sea pequeño, y al hacer esto, destruye la potencialidad del Teorema del Límite Central de garantizar una distribución Normal. Podría ser obvio cual parte de los datos han sido afectados y los puntos anormales (localizados presuntamente muy alejados de la zona crítica de los datos) pueden no ser considerados al hacer el cálculo de σ . Pero a la vez podrí no ser tan obvio. Por ejemplo, en la Fig. 3(a), el punto más intenso está tan alejado de su vecino mas cercano que es claramente el resultado de una cambio de corta duración en la distribución y nadie argumentarí no incluirlos en la estimación de σ . Pero, ¿que hacer con los otros puntos? El rechazar uno ó dos ó más puntos *en forma a posteriori* y sin ninguna justificación física llevaría al experimentador por un camino peligroso; de forma igualmente peligrosa será escoger *una pequeña porción* de los datos para calcular la estimación del ruido. Es recomendable adoptar un método *a priori*, decir exactamente en qué consiste, presentar tantos datos como sea posible (una imagen representa más de mil palabras) y, con un poco de experiencia (negativa) ser pesimista más que optimista. Esto sería suficiente para describir los resultados que uno ha obtenido; para la interpretación de los resultados de otros, parece existir la gran necesidad de conocer al experimentador.

Una *prueba* rigurosa de una *hipótesis* como se describe en la mayoría de los libros en estadística (por ejemplo en las referencias [1] y [4]) no es tan fácil de aplicar a la preguntas de detección de señales. Sin embargo, la esencia de la detección de señales

reside en otro lugar, y es ahí donde existe bastante bibliografía, por ejemplo referencias [6] y [7]. El secreto del éxito consiste en conocer el resultado y entonces filtrar la información anormal para mejorar el cociente de señal a ruido de la respuesta. El "resultado" consiste de las componentes de Fourier en la matriz de datos que produce el sistema de radio telescopio y objeto celeste; y con justicia, algunos de estas posibles componentes pueden ser rechazadas en forma específica si el observador no está interesado en ellas (ó en forma accidental si él ó ella no está consiente de que existan). El proceso de filtrado reduce la amplitud de las componentes de Fourier no deseadas. Siempre resulta en alguna pérdida de información porque las componentes *deseadas* y *no deseadas* deben traslaparse en algún lugar en el dominio de frecuencias. El filtrado puede ser mecánico, electrónico ó digital. Actualmente es práctica común dejar que el instrumento obtenga los datos con un mínimo de interacción ó tratamiento y trabajar con ellos (los datos) en forma digital después.

Considérese por ejemplo los datos de la Fig. 3(a). Contiene ruido (¡ mucho!) y existe un poco de deslizamiento de la señal de comparación ó base y quizás un poco de señal (¡ no lo suficiente!). Es por esto que deseamos suprimir las frecuencias espaciales *más altas*, las cuales son debidas al ruido y son mucho más intensas que aquéllas en la Transformada de Fourier de la respuesta del radio telescopio (con un ancho a potencia media de 4 minutos de arco), y las frecuencias espaciales *más bajas* aquéllas longitudes de onda comparables a las longitud de los datos. Veamos primero lo concerniente al ruido. Para esto necesitamos un filtro digital pasa-bajas el cual consiste simplemente de un conjunto de "pesos" y cada punto "filtrado" se obtiene como el promedio sopesado de sí mismo y un número de sus vecinos en cada lado – en este ejemplo se utilizaron cinco puntos en cada lado; si $x f_i$ son los puntos filtrados y x_i son los puntos originales, se tiene

$$x f_i = \frac{x_i + \sum_{j=1}^5 w_j (x_{i-j} + x_{i+j})}{1 + 2 \sum_{j=1}^5 w_j}.$$

Los pesos $w_j (\leq 1)$, calculados por Martin [8], esaban pensados para dar un filtro muy severo para no tener ninguna señal como sobrepico y con un ancho a media potencia de

aproximadamente un medio de la respuesta del telescopio produciendo un poco de pérdida de resolución. La efectividad del método para el tratamiento del ruido se muestra en la Fig. 3(b). (Un filtro más sencillo que omite el propósito del diseño en su totalidad es sorprendentemente efectivo: simplemente se promedia cada punto con sus vecinos y se omite cualquier pérdida de resolución al escoger un pequeño número de vecinos (¡de forma simétrica!) de tal forma que el ancho de este filtro *cuadrado* (todos los pesos $w_j = 1$) es menor que digamos, un medio del ancho de las líneas ó señales esperadas.)

Ahora bien, ¿es necesario tener un filtro pasa-altas para mejorar la línea de comparación? Para una serie de datos corta tal como se muestra, un ajuste lineal de la línea de base ó de comparación es probablemente satisfactoria: promediar varios puntos en cada lado, ajustar una línea recta, y sustraerla de todos los puntos. Claramente esto cambia la Fig. 3(b) de una manera análoga como si se hubiera rotado la página, y nos dá un poco más de interpretación – una deflexión con la forma del haz ó respuesta del radio telescopio cerca en la posición indicada sugiere la detección del par de galaxias NGC 3309 y NGC 3311, confirmada en forma independiente con observaciones a una frecuencia de 5 GHz utilizando un técnica observacional completamente diferente. Un filtrado pasa-altas que quite las frecuencias más bajas puede ser importante para una serie de datos más larga cuya longitud exceda la escala de las señales individuales por factores muy grandes. El proceso usualmente consiste de la substracción de la línea de comparación (horizontal) ó estimación del continuo y su substracción subsecuente. Un método usualmente utilizado consiste en realizar un ajuste de mínimos cuadrados de un polinomio (por ejemplo ver referencia [9]). En la práctica, esto resulta ser muy tedioso, y *suavizar fuertemente* quizás nos dé un mejor resultado. Esto consiste en construir una matriz de línea de base en la cual cada punto es el promedio del punto correspondiente en la matriz de datos y *muchos* de sus vecinos en cada lado, digamos hasta ± 3 ó ± 4 veces la dimensión esperada de las señales.) Fig. 4 demuestra la efectividad de esta técnica de filtrado pasa-altas de una serie de datos muy larga en radio astronomía. Es efectiva para cualquier serie de datos (de continuo) donde las señales ocupan ≤ 50 por ciento, y en particular, proveé una buena manera de estimar la emisión del continuo en la banda óptica de un espectro digitalizado.

Estas técnicas de ajuste son procesos lineales en la mayoría de los casos y por lo tanto

el cálculo para mejorar el cociente de señal a ruido, los errores apropiados y los niveles de significancia son relativamente simples en lo que se refiere a su interpretación. Toma poca experiencia en este tipo de análisis de datos el descubrir que el ojo humano (junto con su computadora conectada en línea) es impresionante en su habilidad para realizar integraciones y filtrado óptimos con un mínimo de información .

Finalmente, es necesario poner atención en lo que proponemos realizar con los datos – es decir, quitarle las componentes de Fourier que creemos que no queremos conocer para poder así mejorar los cocientes de señal a ruido de aquellos datos que si queremos. Esto no significa para nada el tipo mas extremoso de búsqueda de señal; algunas consisten en tratar de ajustar perfiles preconcebidos a cada punto en la matriz de datos, ó de tratar de ajustar secuencias de perfiles. El tratamiento digital de datos a través de la computadora puede esperarse de dar menores prejuicios (ó sesgos) y ser más objetivo. Quizás sea el caso en el sentido más estrecho; en un sentido más amplio podría producir un sesgo muy extremo. Si fuésemos lo suficientemente tontos para llevar a cabo tal análisis en forma ciega, podemos estar absolutamente seguros que nunca descubriremos nada más allá de lo que sabemos ó de lo que creemos que sabemos.

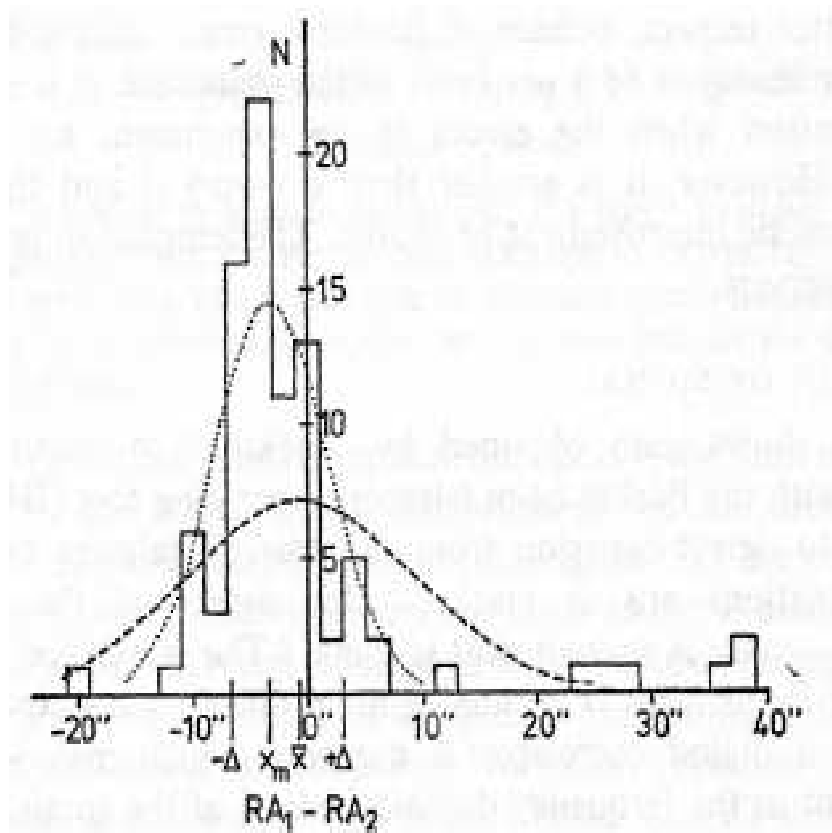


Figure 2: Histograma de las diferencias entre conjuntos ficticios de mediciones de ascensión recta. La curva punteada es la curva Normal "describiendo" la distribución en base a un cálculo directo de \bar{x}_s y σ . Una mejor descripción se obtiene cuando la mediana, x_m se utiliza como localización de la media y $\pm\Delta$ conteniendo 68.3 por ciento de los datos como la desviación estándar. La curva Normal utilizando estos parámetros como medidas de \bar{x} y σ se muestra como la curva punteada.

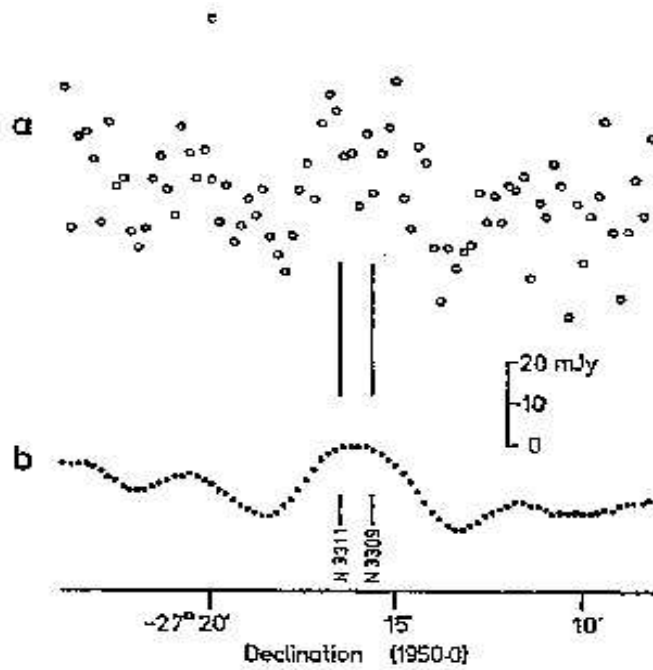


Figure 3: (a) Suma de 50 rastreos en declinación, 2°.5 por minuto, para detectar el par de galaxias NGC 3309 y NGC 3311 con el radio telescopio de 64m de diámetro en Parkes, Australia a una frecuencia de onda de 5 GHz. (b) Las observaciones después de la aplicación de un filtro digital para atenuar las frecuencias espaciales más allá de la respuesta del haz principal del telescopio. Las líneas verticales indican las posiciones en declinación de las galaxias tomadas de imágenes ópticas.

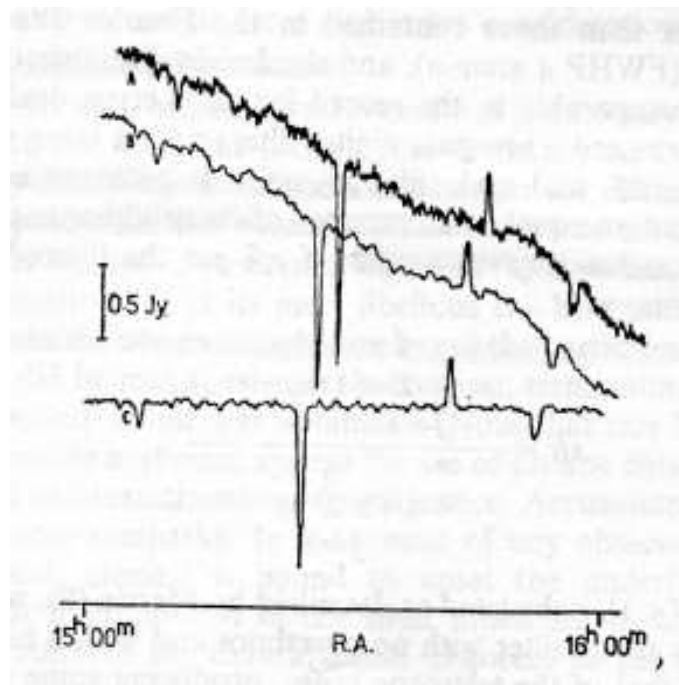


Figure 4: Pasos a seguir en el proceso de análisis de datos (ver referencia[10]) durante las observaciones en Parkes de un estudio extragaláctico a 2.7 GHz. En el estudio se utilizó un sistema de un par de haces principales idénticos en el cual el receptor grafica la diferencia entre las señales entre los dos receptores del radio telescopio de 64m; de ahí la apariencia de fuentes "positivas" y "negativas" en los datos. El rastreo A – tal y como se obtuvo del muestreo de la salida del receptor; el rastreo B – después de aplicar un filtro pasa bajas para atenuar las frecuencias de ruido; rastreo C – después de substraer una curva de comparación (un filtro pasa altas) que consiste de la construcción de una matriz de datos de la línea de base con un alto suavizamiento y la substracción de esta matriz de datos del rastreo B.

Agradecimientos

Deseo agradecer a John Shakeshaft y Richard Hills por amplias conversaciones y críticas asertivas.⁷

⁷JAG-B agradece al Instituto de Astronomía, UNAM por el permiso de ir de visita seis meses (Septiembre 2003 a Febrero 2004) al Departamento de Radio Astronomía Milimétrica del Instituto Tecnológico de California en Pasadena, California, EUA; al Prof. Nick Scoville y al grupo de Radio Astronomía Milimétrica del Instituto Tecnológico de California en Pasadena, California, EUA por su hospitalidad y a DGAPA, UNAM (México) por su apoyo económico parcial que le han permitido realizar esta estancia de investigación, durante la cual se inició esta traducción.

References

- [1] Kendall, M.G., & Stuart, A. *Advanced Theory of Statistics*: (1969) Vol. 1, *Distribution Theory*: (1973) Vol. 2 *Inference and Relationship*: (1976) Vol. 3 *Design and Analysis, and Time Series*, Charles Griffin, London. Un gran texto académico; denso pero con muchos ejemplos.
- [2] Siegel, S., 1956. *Nonparametric Statistics for the Behavioral Sciences*, McGraw Hill Kogakusha, Tokyo. Elemental en el extremo, pero es una gran libro en lo referente a métodos no paramétricos. Sin ninguna teoría o justificación pero con muchas referencias a los textos originales. Todo está ilustrado con ejemplos completamente detallados.
- [3] *The Fontana Dictionary of Modern Thought 1976*, Eds. Bullock, A. & Stallybrass, O., Fontana Books, London. Lleno de cosas interesantes, la mayoría de las cuales no tienen nada que ver (y son aún más fascinantes que) con la estadística.
- [4] Martin, B.R., 1971. *Statistics for Physicists*, Academic Press, London. Muy conciso, fácil de leer y medio riguroso; contiene un gran número de ejemplos con detalle.
- [5] Jauncey, D.L. 1968, *The Astrophysical Journal*, 152, 647.
- [6] Davenport, W.B. & Root, W.L. 1958. *An Introduction to the Theory of Random Signal and Noise*, McGraw Hill, Nueva York.
- [7] Schwartz, M. & Shaw, L., 1975. *Signal Processing: Discrete Spectral Analysis, Detection, and Estimation*, McGraw Hill, Nueva York. Ambos trabajos contienen mucha información sobre la teoría de estadística/información/muestreo, autocorrelación, física del ruido, etc.
- [8] Martin, M.A., 1959. *IRE Trans.Space Electron.Telem. SET-5*, p. 33.
- [9] Bevington, P.R., 1969. *Data Reduction and Error Analysis for the Physical Sciences*, McGraw Hill, Nueva York. Un libro extremadamente útil, elemental, de fácil lectura,

con justificaciones fáciles (pero no rigurosas) y explicaciones; con muchos ejemplos en detalle y muchas subrutinas como programas para computadora (FORTRAN).

[10] Wall, J.V., 1973, *Proc.Astronom.Society Australia*, 2, 195.