Universidad de Alcalá

Estadística-780004

Unidad Temática 1. Estadística Descriptiva

Tema 1. Descripción de una variable

Juan J. Cuadrado Gallego

T1. Descripción de una variable

El número de datos a describir de una característica cuantitativa variable puede ser muy grande. Para facilitar su análisis estadístico se pueden agrupar los datos utilizando Clases de Equivalencia.

Para agrupar los datos en clases de equivalencia se deben realizar los siguientes cálculos:

1. Rango. Con los datos ordenados por magnitud, se resta el menor del mayor.
2. Número de clases. Dependerá del número de datos.
3. Límites. Son los números menor y mayor se cada clase.
4. Amplitud. Límite superior menos límite inferior de una clase.
5. Marca. Punto medio. Representa a la clase. Suma de los límites de la clase dividido entre 2.

T1. Descripción de una variable

En el primer análisis de los datos se cuantifica la frecuencia de aparición de los mismos. [1] define Frecuencia como: 3.f.Estad. Número de elementos comprendidos dentro de un intervalo en una distribución determinada.

En Estadística dicha definición se amplia y profundiza a través de los siguientes conceptos:

Frecuencia Absoluta f_i: Número de apariciones de un dato dado.

Frecuencia Relativa f_{ri}: Número de apariciones de un dato dado dividido entre el número de datos.

Frecuencia Acumulada (Absoluta o Relativa): Con los datos ordenados por magnitud, suma de las frecuencias absolutas o relativas de los datos inferiores al dato más la del dato.

Distribución de Frecuencias (Absoluta o Relativa): Pares compuestos de cada dato y su frecuencias.

Para datos agrupados en clases de equivalencia, el concepto de dato se cambia por el de clase.

T1. Descripción de una variable

El segundo análisis de los datos se basa en el cálculo de la media y la moda.

[1] define Media como: 3. Mat. Número que resulta al efectuar una serie determinada de operaciones con un conjunto de números y que, en determinadas condiciones, puede representar por si solo a todo el conjunto. Recibe diferentes denominaciones según las operaciones que se realicen para obtenerlo, así: media aritmética, media geométrica, etc.

La media más utilizada en Estadística es la Aritmética:

$$\bar{x}_a = \frac{\sum_{i=1}^{n} f_i x_i}{F_n} \quad \text{donde} \quad F_n = \sum_{i=1}^{n} f_i$$

Se denomina media aritmética ponderada: Si en lugar de f_i se consideran los pesos w_i

Existen otras medias que se deben aplicar en determinados, por ejemplo la Geométrica:

$$\bar{x}_g = \left(\prod_{i=1}^{n} x_i^{f_i}\right)^{1/F_n}$$

Para datos agrupados en clases de equivalencia, el concepto de dato se cambia por el de clase.

T1. Descripción de una variable

Juan J. Cuadrado Gallego

Moda (x_{MOD}): Es el valor más frecuente.
La moda no siempre es única. Si la población o la muestra tiene dos modas es *bimodal*, si tiene tres, *trimodal*, etc.

T1. Descripción de una variable

El tercer análisis de los datos se basa en el cálculo de las medidas de dispersión. [1] define Dispersión como: 4. f. Mat. Distribución estadística de un conjunto de valores.

Las medidas de dispersión más utilizadas en Estadística son:

1. Dispersión Absoluta. Se calcula a través de las medidas Desviación, Varianza y Rango. La dispersión absoluta puede calcularse sobre las media o sobre las medidas de ordenación.

2. Dispersión Relativa. Cociente entre la media aritmética y la dispersión absoluta.

   Si la dispersión absoluta se ha calculado mediante la desviación típica, la dispersión relativa se denomina Coeficiente de Variación de Pearson y se calcula como:

$$CV = \frac{s}{\bar{x}}$$

T1. Descripción de una variable

[1] define Desviación como: Diferencia entre la medida de una magnitud y el valor de referencia.

Las desviaciones más utilizadas en Estadística son:

Desviación Estándar o Típica:

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{\sum_{i=1}^{n} f_i}}$$

Desviación Media:

$$s = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{\sum_{i=1}^{n} f_i}$$

Desigualdad de Tchebychev: En cualquier distribución entre la media y k desviaciones típicas se encuentran, al menos, el $\left(1 - \frac{1}{k^2}\right).100\%$ de los datos.

Para datos agrupados en clases de equivalencia, el concepto de dato se cambia por el de clase.

Juan J. Cuadrado Gallego

T1. Descripción de una variable

[1] define Varianza como: 1.f.Estad. Media de las desviaciones cuadráticas de una variable aleatoria, referidas al valor medio de esta.

Varianza:

$$s^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{\sum_{i=1}^{n} f_i}$$

Para datos agrupados en clases de equivalencia, el concepto de dato se cambia por el de clase.

T1. Descripción de una variable

En el cuarto análisis de los datos se basa en las medidas de ordenación.

[1] define Mediana ($\tilde{x}$) como: 9.f. Mat. Elemento de una serie ordenada de valores crecientes de forma que la divide en dos partes iguales, superiores e inferiores a él

La mediana se calcula de la siguiente manera:

1. Se ordenan los datos por magnitud.
2. Para un número par de datos $\tilde{x}$ es la suma de los dos valores centrales dividido entre 2:

$$\tilde{x} = \frac{x_{n/2} + x_{(n/2)+1}}{2}$$

3. Para un número impar de datos $\tilde{x}$ es el valor central:

$$\tilde{x} = x_{(n+1)/2}$$

T1. Descripción de una variable

Cuantiles: Los cuantiles permiten dividir el conjunto ordenado de datos en un conjunto de partes de igual tamaño.

Los más utilizados son los Percentiles, Cuartiles y Deciles:

- [1] define Percentil ($\tilde{x}_p;\ p = \frac{1}{100}, \frac{2}{100}, \ldots \frac{99}{100}$) como: 1. m. Mat. Valor que divide un conjunto ordenado de datos estadísticos de forma que un porcentaje de tales datos sea inferior a dicho valor. Así, un individuo en el percentil 80 está por encima del 80% del grupo a que pertenece.
- y define Cuartil ($\tilde{x}_c;\ c = \frac{1}{4}, \frac{2}{4}, \frac{3}{4}$) como: m. Mat. Cualquiera de los percentiles 25, 50 ó 75
- Un Decil ($\tilde{x}_d;\ d = \frac{1}{10}, \frac{2}{10}, \ldots \frac{9}{10}$) es cualquiera de los percentiles 10, 20,…50,…, 90, 100

T1. Descripción de una variable

Para datos no agrupados, los cuantiles se calculan como (para *c* y *d* sería igual cambiando *p*):

$\tilde{x}_{p,c,d} = x_{[np]+1}$ si $np \notin \mathbb{N}$ y

$\tilde{x}_{p,c,d} = \frac{x_{np}+x_{np+1}}{2}$ si $np \in \mathbb{N}$

$[np]$ parte entera

T1. Descripción de una variable

Juan J. Cuadrado Gallego

Medidas de dispersión basadas en medidas de ordenación.

[1] define Rango como: 5.m.Estad. Amplitud de la variación de un fenómeno entre un límite menor y uno mayor claramente especificados. Los rangos más utilizados en estadística son:

- Rango Intercuartílico: $Rc = \tilde{x}_{3/4} - \tilde{x}_{1/4}$
- Rango Interpercentilico: $Rp = \tilde{x}_{90/100} - \tilde{x}_{10/100}$
- Rango Interdecilico: $Rd = \tilde{x}_{9/10} - \tilde{x}_{1/10}$

Desviación:

- Desviación Intercuartílica (Rango semiintercuartilico): $Dc = \frac{\tilde{x}_{3/4} - \tilde{x}_{1/4}}{2}$
- Desviación Interpercentilica: $Dp = \frac{\tilde{x}_{90/100} - \tilde{x}_{10/100}}{2}$
- Desviación Interdecilica: $Dd = \frac{\tilde{x}_{9/10} - \tilde{x}_{1/10}}{2}$

T1. Descripción de una variable

Los datos también pueden ser analizados estadísticamente mediante técnicas que permiten su visualización. [1] define Visualizar como: Formar en la mente una imagen visual de un concepto abstracto.

Las dos técnicas más utilizadas son las Tablas y Gráficas (que se llamarán diagramas):

- [1] define Tabla como: 2. f. Lista o catálogo de cosas puestas por orden sucesivo o relacionadas entre sí.
- [1] define Gráfica como: 4. m. Representación de datos numéricos por medio de una o varias líneas que hacen visible la relación que esos datos guardan entre sí.

T1. Descripción de una variable

Algunos Formatos de Tablas utilizadas

x	f	f_r
x_1	f_1	f_{r1}
x_k	f_k	f_{rk}
x_n	f_n	f_{rn}
	F_n	1

a)

x	x_1	x_k	x_n	
f	f_1	f_k	f_n	F_n
f_r	f_{r1}	f_{rk}	f_{rn}	1

b)

$l_{i-1} - l_i$	m	f	f_r	f_A	f_{rA}	a_i	h_i

c)

Iguales para datos cualitativos y cuantitativos pero los cualitativos no tienen acumulados

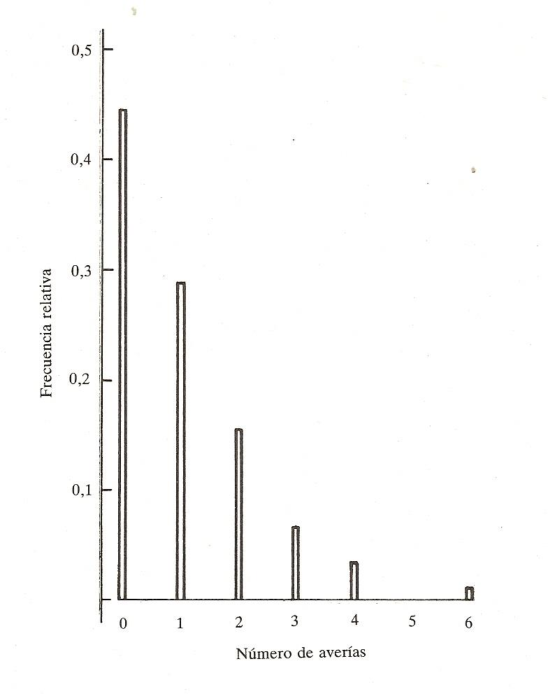
Juan J. Cuadrado Gallego

T1. Descripción de una variable

Diagrama de Barras

Se representan datos cualitativos y cuantitativos discretos.

Cada característica se representa por un rectángulo cuya altura es su frecuencia absoluta o relativa.



T1. Descripción de una variable

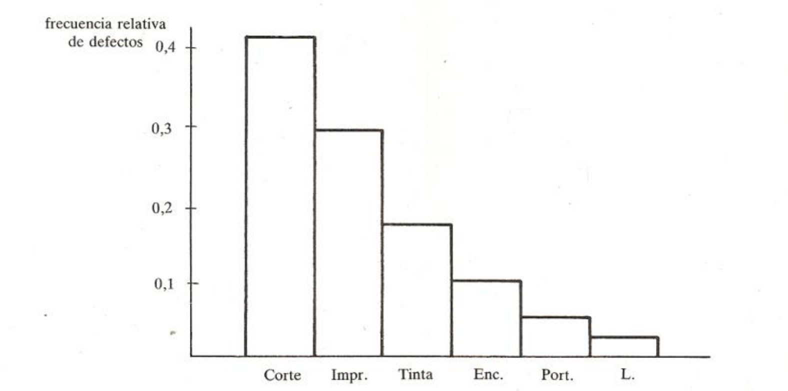
Juan J. Cuadrado Gallego

Diagrama de Frecuencias de Pareto.

Se representan datos cualitativos.

Se ordenan las clases por su f_r

Cada característica se representa por un rectángulo cuya altura es su frecuencia relativa o absoluta



Ley de Pareto: Regla 80-20. El 80% de los efectos se encuentra en el 20% de las causas, el 80% de los fallos corresponden al 20% del código.

T1. Descripción de una variable

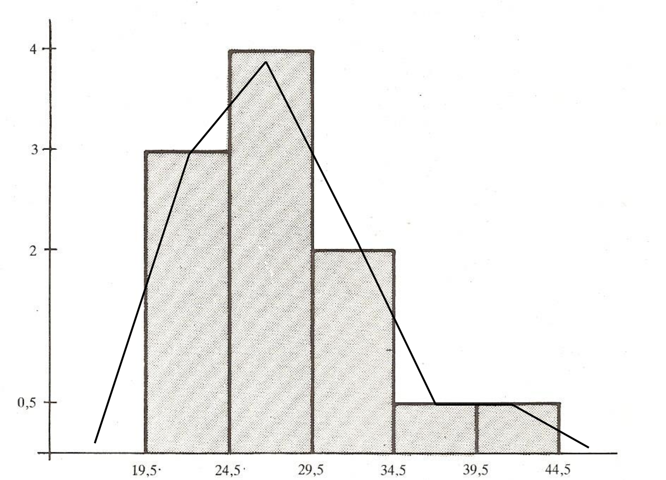
Juan J. Cuadrado Gallego

Diagrama de Frecuencias Histograma.

Se representan datos cuantitativos continuos agrupados.

Cada rectángulo representa una clase cuya base es proporcional a la amplitud del intervalo.

La altura se determina de manera que su área sea proporcional a la frecuencia de la clase. Si la amplitud es constante la altura coincide con la frecuencia. Si la amplitud es variable la altura se calcula como: $h_i = \frac{n_i}{a_i}$ (Área de un rectángulo)



El Polígono de Frecuencias se construye a partir del Histograma o el Diagrama de Barras uniendo los puntos superiores de las barras o rectángulos (medios) y el eje horizontal al principio y al final

T1. Descripción de una variable

Diagrama de Tallo y Hojas de Tukey.

Se representan datos cuantitativos. Es como un histograma con barras no vacías sino formadas por los datos. Se deben redondear los datos.

Trazar una línea vertical. Se representa a la izquierda el dígito/s de la clase (tallo) y a la derecha el de las unidades (hojas). EL tallo define la clase y las hojas su frecuencia.

Juan J. Cuadrado Gallego

T1. Descripción de una variable

Diagrama de Caja y Bigotes de Tukey.

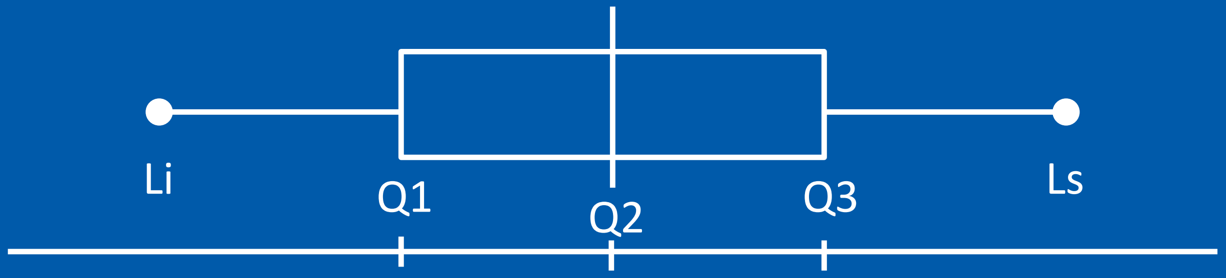
Se representan datos cuantitativos continuos.

Se Ordenan los datos y se obtiene el valor mínimo, el máximo y los cuartiles.

Se dibuja un rectángulo con extremos en Q1 y Q3.

Se Indica la posición de la mediana Q2 con una línea.

Se calculan los límites del intervalo para los valores atípicos y se conectan con una línea al rectángulo



Permite identificar muy fácilmente los valores atípicos como aquellos que están fuera del intervalo

Juan J. Cuadrado Gallego

T1. Descripción de una variable

Los datos atípicos o outliers son observaciones muy diferentes del resto. La identificación y análisis de los Datos Atípicos es fundamental para realizar un estudio estadístico correcto.

Los outliers son:

- O bien datos erróneos procedentes de errores de medida o transcripción, que deben ser eliminados ya que alteran erróneamente las medidas obtenidas sobre los datos analizados.
- O bien son datos correctos con mucha significación ya que se apartan de lo normal y pueden conducir a hallazgos importantes.

Existen diferentes maneras de identificarlos y diferentes teorías sobre que datos deben ser considerados atípicos.

Universidad de Alcalá

Juan J. Cuadrado Gallego

Estadística-780004

Unidad Temática 1. Estadística Descriptiva

Tema 2. Descripción conjunta de varias variables

T2. Descripción conjunta de varias variables

Análisis de dos características

El análisis conjunto de dos características, x e y, se realiza simultáneamente sobre dos conjuntos de datos, correspondientes, cada uno de ellos, a cada una de las características estudiadas.

Los tipos de datos que se pueden recoger sobre las características estudiadas son:

- x e y cualitativas
- x cualitativa e y medible, o a inversa
- x e y medibles

T2. Descripción conjunta de varias variables

Frecuencia

Frecuencia Absoluta f_{ij}: Número de apariciones del par x_i, y_j.

Frecuencia Relativa f_{ri}: Número de apariciones del par x_i, y_j. dividido entre el número de pares de datos $n.m$.

$$f(x_i, y_j) \qquad f_r(x_i, y_j) = \frac{f(x_i, y_j)}{n.m} \qquad \sum_{i=1}^{n} \sum_{j=1}^{m} f_r(x_i, y_j) = 1$$

T2. Descripción conjunta de varias variables

Frecuencia

Distribución de frecuencias de dos variables. Pares de datos y sus frecuencias absolutas o relativas.

Tabla de doble entrada: Representa los valores de ambas variables y sus frecuencias. Si las variables son cualitativas se denomina Tabla de Contingencia.

Distribución marginal de una variable: valores de las frecuencias de la variable.

$$f_r(x_i) = \sum_{j=1}^{m} f_r(x_i, y_j) \qquad f_r(y_j) = \sum_{i=1}^{m} f_r(x_i, y_j)$$

T2. Descripción conjunta de varias variables

Juan J. Cuadrado Gallego

Las medidas de dependencia más utilizadas en estadística son la Covarianza y la Correlación.

La Covarianza se define como:

La media aritmética de los productos de las desviaciones de cada variable respecto a su media aritmética o la media de los productos menos el producto de las medias.

$$s_{xy} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}\,(x_i - \bar{x})(y_j - \bar{y})}{\sum_{i=1}^{n} f_i}$$

$$= \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}\,x_i y_j}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}\right)\cdot\left(\frac{\sum_{j=1}^{m} f_j y_j}{\sum_{j=1}^{m} f_j}\right)$$

T2. Descripción conjunta de varias variables

Se define la Matriz de Covarianzas de las variables como una matriz cuadrada simétrica cuya diagonal principal está formada por las varianzas de los valores observados y la diagonal opuesta por las covarianzas. En el caso de dos variables la matriz toma la forma:

$$\bar{M} = \begin{bmatrix} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{bmatrix}$$

T2. Descripción conjunta de varias variables

Las medidas de dependencia más utilizadas en estadística son la Covarianza y la Correlación.

[1] define Correlación como: 4. f. Mat. Medida de la dependencia existente entre variantes aleatorias.

La correlación lineal se calcula a través de la ecuación:

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

r toma valores entre -1 y 1.

- Si la correlación lineal es perfecta, es decir si los valores de *x* e *y* están sobre una recta el valor de *r* será -1 si tiene pendiente negativa y 1 si la tiene positiva.
- Si *r* es igual a 0 no hay dependencia lineal entre las variables, lo cual implica o bien que las variables son independientes, o bien que hay una dependencia no lineal entre las mismas.

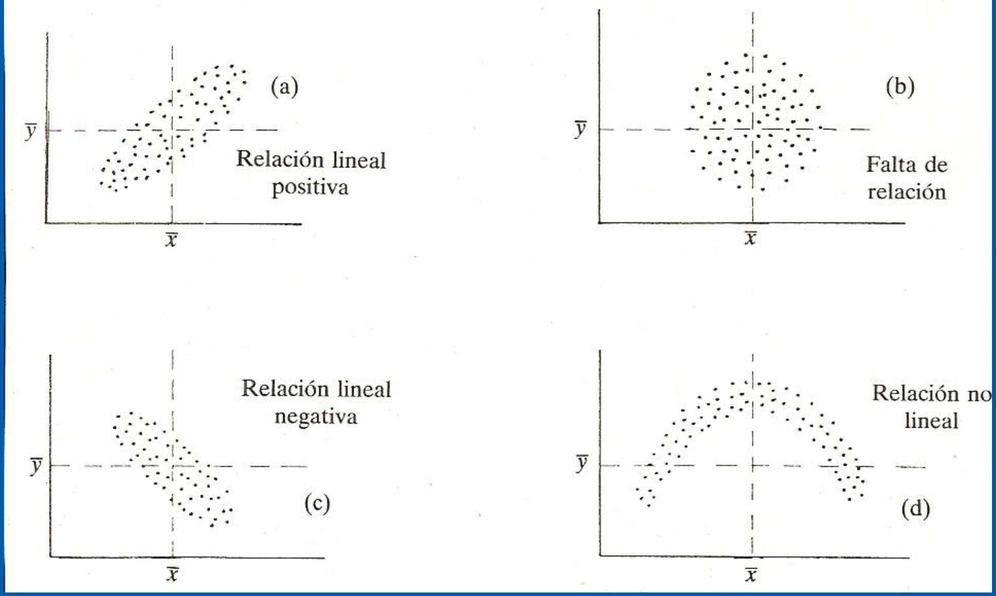
T2. Descripción conjunta de varias variables

Juan J. Cuadrado Gallego

Las relación entre dos o más variables se establece a través de funciones.

[1] define Relación (Regression) como: . f. Mat. Resultado de comparar dos cantidades expresadas en números.

El primer paso en el análisis de regresión consistirá en la visualización de los pares de datos mediante un gráfico denominado diagramas de dispersión o nube de puntos



T2. Descripción conjunta de varias variables

El segundo paso consistirá en la obtención de la función de función de regresión.

Una vez establecida la función de regresión podremos obtener el valor de una de las variables si conocemos el valor de la otra.

Las relación entre dos variables más básica es una relación lineal.

La función de relación lineal entre dos variables es una recta.

$$y = a + bx$$

Se deben determinar a y b

T2. Descripción conjunta de varias variables

El método de ajuste más utilizado es el de mínimos cuadrados, que está basado en encontrar una función para calcular la *y* a partir de *x*, tal que minimice el valor de la suma del cuadrado las restas de todos los valores de *y* observados menos los calculados.

$$\sum_{i=1}^{n} (y_i - y_{ci})^2 \text{ Mínimo}$$

Los valores de *a* y *b* a partir del método de mínimos cuadrados son:

$$b = \frac{s_{xy}}{s_x^2} = r_{xy}\frac{s_y}{s_x} \qquad \mathrm{a} = \bar{y} - b\bar{x}$$

T2. Descripción conjunta de varias variables

Existen diferentes análisis que indican en qué medida una recta establece correctamente la relación entre dos variables *x* e *y*. El primero que vamos a estudiar es el ANOVA. Para medirlo calculamos:

La dispersión de los valores de y calculados a través de la función de regresión, $\hat{y}_i$. Este valor se calcula como la suma de los cuadrados de la diferencia entre el valor de calculado para cada x a través de la ecuación de la recta, $\hat{y}_i$ y la media del valor de y, $\bar{y}$.

$$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

La dispersión de los valores de y observados, y_i. Este valor se calcula como la suma de los cuadrados de la diferencia entre el valor observado para cada y y la media del valor de y, $\bar{y}$.

$$SSy = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Una vez obtenidos los dos valores anteriores, se calcula la correlación cuadrada r^2 como la razón entre SS$\hat{y}$ y SSy. La ecuación de cálculo es:

$$r^2 = \frac{SSR}{SSy}$$

r^2 tendrá un valor que variará entre 0 y 1. De tal forma que 0 indica que no hay ajuste entre ambas magnitudes, mientras que 1 indica que el ajuste es perfecto.

T2. Descripción conjunta de varias variables

Existen diferentes análisis que indican en qué medida una recta establece correctamente la relación entre dos variables *x* e *y*. El segundo que vamos a estudiar es el error estándar de la estimación o desviación típica residual. Para medirlo calculamos:

$$s_r = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_{ci})^2}{n}}$$

Mide la dispersión de los valores (x, y) observados alrededor de los valores (x, y) calculados a través de la recta de ajuste. Cuanto más cercano esté a 0 más cerca estarán los valores de la recta. Es similar a la desviación típica con respecto a la media aritmética para una variable.

Sobre el 95% de los puntos están situados en la región comprendida entre dos líneas paralelas a la recta de regresión separadas por 4 veces s_r, 2 veces a cada lado de la recta. Y sobre 66% de los datos en una separación de 2 veces s_r, 1 vez a cada lado de la recta

Universidad de Alcalá

Juan J. Cuadrado Gallego

Estadística-780004

Unidad Temática 2. Probabilidad

Tema 3. Probabilidad

T3. Probabilidad

¿Por qué se estudia probabilidad al estudiar estadística?

Cuando los datos que se estudian son una muestra de una población, el objetivo del estudio es inducir (estadística inductiva) las propiedades de la población a partir de la muestra.

Los modelos estadísticos, basados en el cálculo de probabilidades, actúan de puente entre lo observado (muestra) y lo desconocido (población)

Por esta razón es imprescindible para poder estudiar estadística inductiva conocer probabilidad

Recordamos las definiciones de Población y Muestra que veíamos en el Tema 0:

[1] define Población como: 5. f. Sociol. Conjunto de los individuos o cosas sometido a una evaluación estadística. (Deseamos conocer alguna/s característica/s)(Puede ser finita o infinita, pero debe ser observable al menos a través de una muestra)

y define Muestra como: 2. f. Parte o porción extraída de un conjunto por métodos que permiten considerarla como representativa de él

T3. Probabilidad

[1] define Experimento como: hacer operaciones destinadas a descubrir, comprobar o demostrar determinados fenómenos o principios. Los experimentos pueden ser:

- Aleatorios: Son aquellos en los que el resultado no está determinado por las condiciones iniciales y es imprevisto.
- Deterministas: Son aquellos en los que el resultado está completamente determinado por las condiciones iniciales.

[1] define Suceso como: En un experimento aleatorio, subconjunto del total de resultados posibles.

A partir de la definición de suceso pueden definirse el concepto de Suceso Elemental: Es cada uno de los resultados más simples que pueden darse en la realización de un experimento aleatorio.

T3. Probabilidad

A partir de la definición de suceso elemental se pueden comenzar a aplicarse al estudio de los sucesos asociados a un experimento aleatorio las matemáticas de la Teoría de Conjuntos, obteniéndose las siguientes definiciones y propiedades:

Espacio Muestral: Conjunto cuyos elementos son los sucesos elementales de un experimento aleatorio. El espacio muestral permite definir el Suceso Seguro: Aquel siempre ocurre.

Suceso Complementario: De un suceso dado A es el que se verifica siempre que no se verifica A.

Suceso Imposible: suceso tal que siempre que se realiza el experimento nunca se obtiene ninguno de los sucesos elementales que lo componen. Es el conjunto vacío (Ø)

Partes de E: conjunto cuyos elementos son todos los posibles subconjuntos de E.

T3. Probabilidad

Juan J. Cuadrado Gallego

[1] define Probabilidad como:

3. f. Mat. En un proceso aleatorio, razón entre el número de casos favorables y el número de casos posibles.

La definición clásica de Probabilidad o Regla de Laplace define la Probabilidad como:

La probabilidad de que aparezca un determinado suceso A, es el cociente entre el número de casos favorables a ese suceso y el número total de casos.

$$P(A) = \frac{n_A}{N}$$

T3. Probabilidad

Las propiedades de una Probabilidad son análogas a las de la frecuencia relativa y son:

1. $0 \leq P(A) \leq 1$
2. $P(\bar{A}) = 1 - P(A)$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. Si $\mathrm{A} \subset B = P(A) \leq P(B)$
5. $P(E) = 1$
6. $P(\emptyset) = 0$

T3. Probabilidad

Desde el punto de vista de como están relacionados entre si, los sucesos pueden ser:

- Independientes: Un conjunto de sucesos son independientes si la ocurrencia de cualquiera de ellos no modifica la probabilidad de ocurrencia del resto.
- Dependientes: Un conjunto de sucesos son dependientes si la ocurrencia de cualquiera de ellos modifica la probabilidad de ocurrencia del resto.

La probabilidad de sucesos independientes viene dada por la ecuación (para dos sucesos):

$$P(A \cap B) = P(A)P(B)$$

La probabilidad de sucesos dependientes se denomina *condicionada* y viene dada por la ecuación (para dos sucesos):

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Donde $P(A|B)$ es la probabilidad de que ocurra A habiendo ocurrido B

T3. Probabilidad

Para obtener el número total de casos *n* se deben contar todos los casos posibles. Esto es sencillo si se trabaja con conjuntos pequeños, pero cuando se trabaja con conjuntos grandes se vuelve muy complicado.

Para poder resolver este problema, se utiliza la combinatoria

[1] define Combinatoria como:

3. f. Mat. Parte de las matemáticas que estudia el número de posibilidades de ordenación, selección e intercambio de los elementos de un conjunto, es decir, las combinaciones, variaciones y permutaciones.

T3. Probabilidad

[1] define Combinación como:

8\. f. Mat. Cada uno de los subconjuntos de un número determinado de elementos de un conjunto finito dado, que difieren al menos en un elemento; p. ej., abc, agc, bcd, acd.

[1] define Variación como:

2\. f. Mat. Cada uno de los subconjuntos del mismo número de elementos de un conjunto dado, que difieren entre sí por algún elemento o por el orden de estos.

[1] define Permutación como:

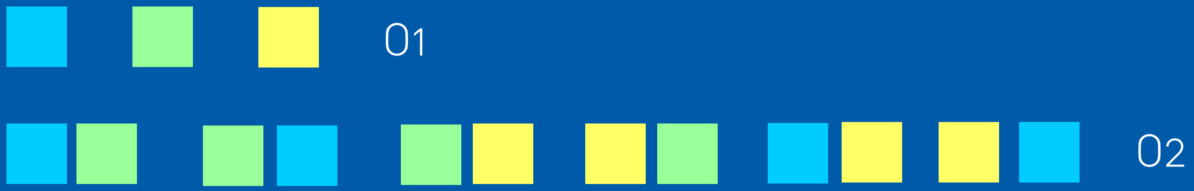
2\. f. Mat. Cada una de las ordenaciones posibles de los elementos de un conjunto finito.

T3. Probabilidad

Juan J. Cuadrado Gallego

Variación:

$$V_m^n = \frac{m!}{(m-n)!}$$

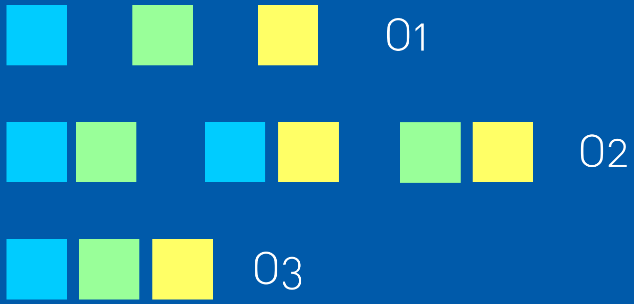


Permutación:

$$P_n = n!$$

Combinación:

$$C_m^n = \frac{m!}{n!(m-n)!}$$

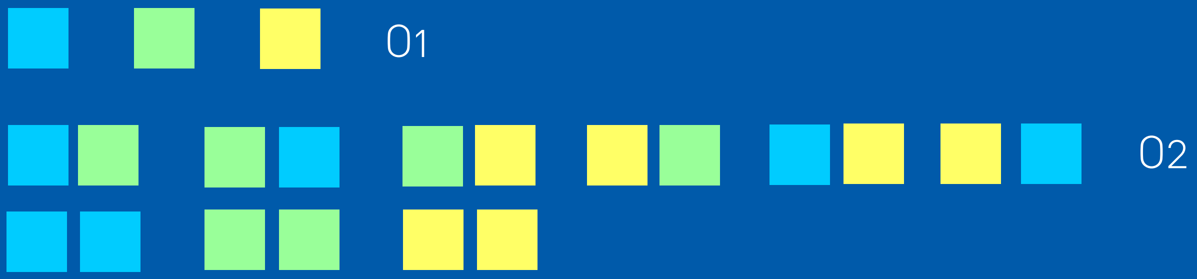


T3. Probabilidad

Juan J. Cuadrado Gallego

Variación con repetición:

$$VR_m^n = m^n$$



Permutación con repetición:

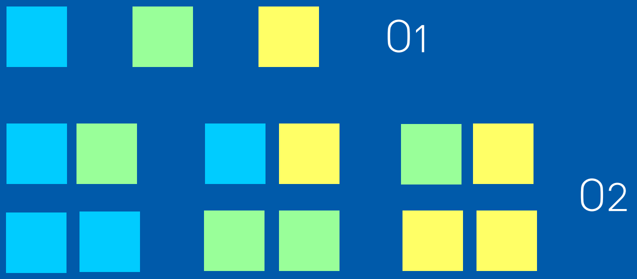
de *n* elementos de los cuales n_1 son iguales, n_2 son iguales pero distintos de los anteriores y así *k* veces, de tal forma que

$$n_1 + n_1 + \cdots + n_k = n$$

$$PR_n^{n_1,n_2,\ldots,n_k} = \frac{n!}{n_1!\,n_2!\ldots n_k!}$$

Combinación con repetición:

$$CR_m^n = \binom{m+n-1}{n}$$



Universidad de Alcalá

Juan J. Cuadrado Gallego

Estadística-780004

Unidad Temática 3. Inferencia Estadística

Tema 4. Variables Aleatorias

T4. Variables Aleatorias

Estudio de la Población a partir de la Distribución de Probabilidad de las Variables

No se conocen los valores exactos de las variables, como es estadística descriptiva, pero, a través del estudio de una muestra, se han determinado sus probabilidades de aparición.

[1] define variable (aleatoria) como: estadística. 1. f. Mat. Magnitud cuyos valores están determinados por las leyes de probabilidad, como los puntos resultantes de la tirada de un dado.

En todo proceso de observación o experimento se puede definir una variable aleatoria asignando a cada resultado del experimento un número:

- Si el resultado es numérico porque se cuenta o mide el valor
- Si el resultado es cualitativo, porque se le hace corresponder un número arbitrariamente

T4. Variables Aleatorias

Como se vio en Estadística Descriptiva, las variables pueden ser discretas y continuas.

Función de Probabilidad: Función matemática discreta que especifica la probabilidad de aparición cada valor del espacio muestral de una variable discreta.

$p(x_i) = P(x = x_i)$ Si E es el espacio muestral $\rightarrow \sum_{i \in E} p(x_i) = 1$

Función de Densidad: Función matemática continua que especifica la probabilidad de aparición cada valor del espacio muestral de una variable continua.

$p(x_i) = f(x = x_i)$ Como f es continua $\rightarrow \int_{-\infty}^{\infty} f(x)dx = 1$

Función de Distribución: Función matemática discreta o continua que especifica para cada valor del espacio muestral de una variable discreta o continua la probabilidad de aparición un valor menor o igual.

$$F(x_i) = P(x \leq x_i)$$

- Si es una variable discreta: $F(x_k) = P(x \leq x_k) = \sum_{i=1}^{k} p(x_i)$ con n valores $\sum_{i=1}^{k} p(x_i) = 1$
- Si es una variable continua: $F(x_k) = P(x \leq x_k) = \int_{-\infty}^{x_k} f(x)dx$

T1. Descripción de una variable (Recordatorio)

En el primer análisis de los datos se cuantifica la frecuencia de aparición de los mismos.

[1] define Frecuencia como:

3.f.Estad. Número de elementos comprendidos dentro de un intervalo en una distribución determinada.

En Estadística dicha definición se amplia y profundiza a través de los siguientes conceptos:

Frecuencia Absoluta f_i: Número de apariciones de un dato dado.

Frecuencia Relativa f_{ri}: Número de apariciones de un dato dado dividido entre el número de datos.

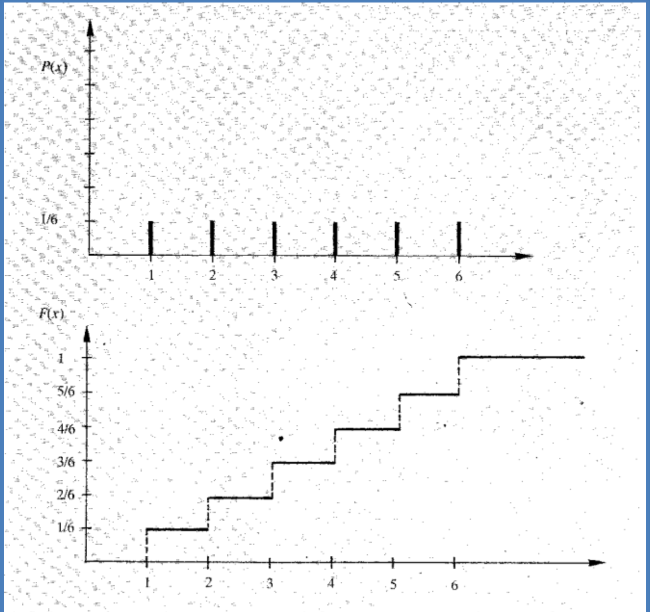
Frecuencia Acumulada (Absoluta o Relativa): Con los datos ordenados por magnitud, suma de las frecuencias absolutas o relativas de los datos inferiores al dato más la del dato.

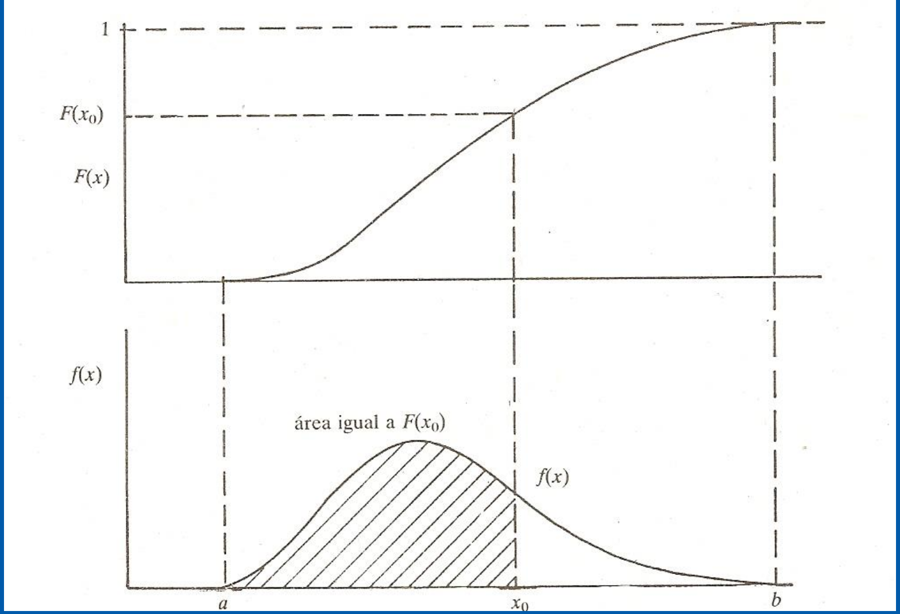
Distribución de Frecuencias (Absoluta o Relativa): Pares compuestos de cada dato y su frecuencias.

Para datos agrupados en clases de equivalencia, el concepto de dato se cambia por el de clase.

T4. Variables Aleatorias

Juan J. Cuadrado Gallego





T4. Variables Aleatorias

Cuando se tiene una variable Discreta o Continua, cuyos valores se pueden inferir con un grado de probabilidad mediante una Función de Probabilidad o una Función de Densidad, respectivamente, se pueden inferir también los mismos parámetros de caracterización de la población que se obtenían para caracterizar poblaciones en Estadística Descriptiva. Éstos son la media (aquí denominada Esperanza) y la varianza.

Esperanza (o media):

- Variable Discreta: $\mu_x = E(x) = \sum_{i=1}^{n} x_i . p(x_i)$
- Variable Continua: $\mu_x = E(x) = \int_{-\infty}^{\infty} x . f(x) dx$

Varianza:

- Variable Discreta: $\sigma_x^2 = \sum_{i=1}^{\infty} (x_i - \mu)^2 . p(x_i)$
- Variable Continua: $\sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu)^2 . f(x) dx = \int_{-\infty}^{\infty} x^2 . f(x) dx - \mu_x^2$

Universidad de Alcalá

Juan J. Cuadrado Gallego

Estadística-780004

Unidad Temática 3. Inferencia Estadística

Tema 4. Modelos Univariantes de Distribución

T4. Modelos Univariantes de Distribución

Como se vio en Estadística Descriptiva, las variables pueden ser discretas y continuas.

Función de Probabilidad: Función matemática discreta que especifica la probabilidad de aparición cada valor del espacio muestral de una variable discreta.

$p(x_i) = P(x = x_i)$ Si E es el espacio muestral $\rightarrow \sum_{i \in E} p(x_i) = 1$

Función de Probabilidad de Bernuilli: Clasifica los elementos de la población dos categorías a las que asigna los valores $0, 1$ $(0 = Correcto)$. Las probabilidades de aparición de los dos valores son constantes.

$p(x_i) = P(x = x_i) = p^x q^{1-x}; con\ x = 0, 1$ // $\mu_x = p$ // $\sigma_x^2 = pq$ // $q = 1 - p;\ p \equiv fallo$

Tiene, al menos, dos funciones asociadas: Binomial y Geométrica:

Se denominan funciones derivadas de Bernuilli porque para la construcción de las correspondientes variables se requiere al menos una función de Bernuilli.

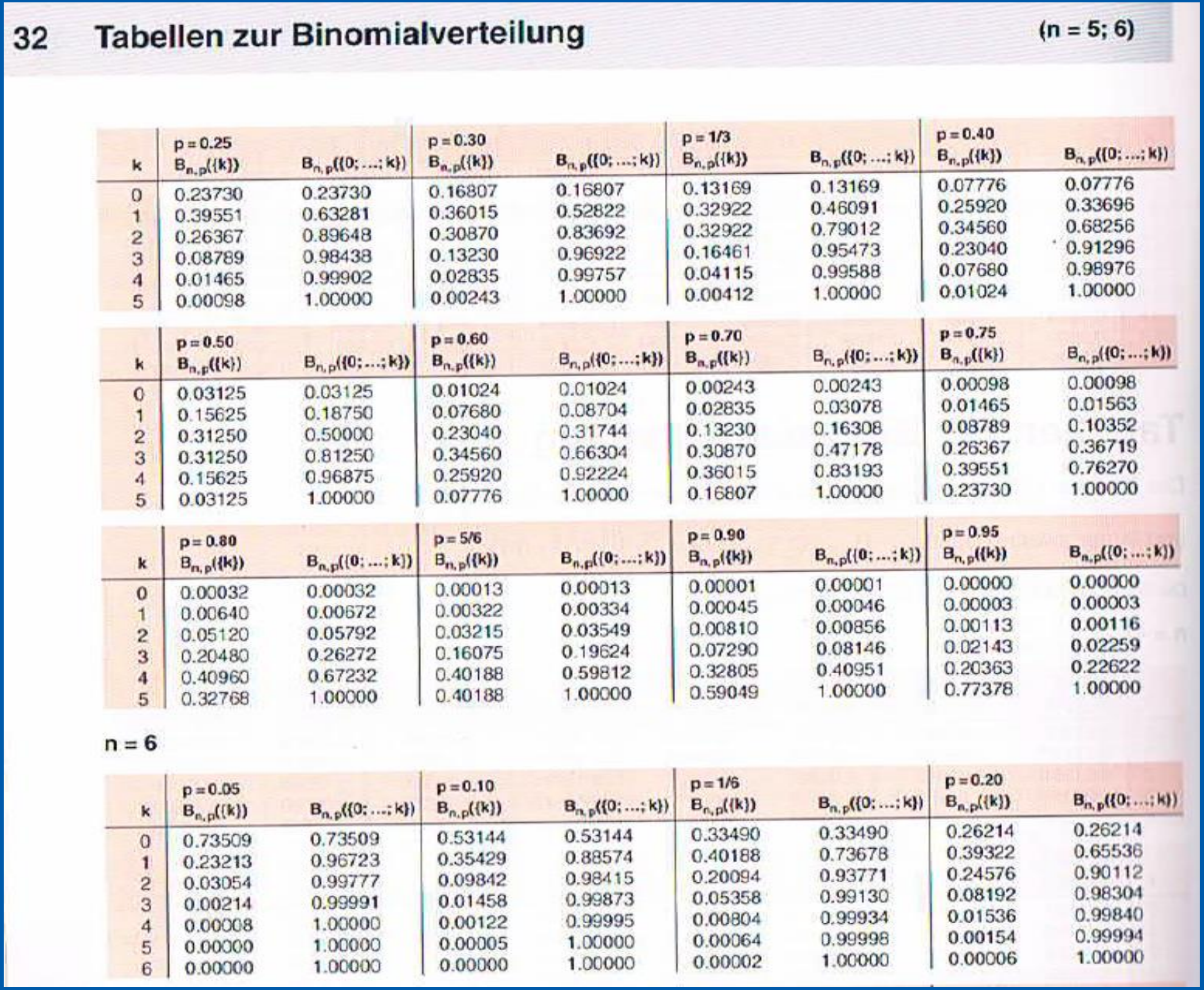
T4. Modelos Univariantes de Distribución

Función de Probabilidad de Binomial: Número de elementos defectuosos al observar *n*.

$$p(r) = P(x = r) = \binom{n}{r} p^r q^{n-r}; con\, r = 0, 1, \dots, n \;\; // \;\; \mu_x = np \;\; // \;\; \sigma_x^2 = npq$$

T4. Modelos Univariantes de Distribución

La Función de Distribución Binomial está tabulada



32 **Tabellen zur Binomialverteilung** (n = 5; 6)

k	p = 0.25		p = 0.30		p = 1/3		p = 0.40	
	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$
0	0.23730	0.23730	0.16807	0.16807	0.13169	0.13169	0.07776	0.07776
1	0.39551	0.63281	0.36015	0.52822	0.32922	0.46091	0.25920	0.33696
2	0.26367	0.89648	0.30870	0.83692	0.32922	0.79012	0.34560	0.68256
3	0.08789	0.98438	0.13230	0.96922	0.16461	0.95473	0.23040	0.91296
4	0.01465	0.99902	0.02835	0.99757	0.04115	0.99588	0.07680	0.98976
5	0.00098	1.00000	0.00243	1.00000	0.00412	1.00000	0.01024	1.00000

k	p = 0.50		p = 0.60		p = 0.70		p = 0.75	
	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$
0	0.03125	0.03125	0.01024	0.01024	0.00243	0.00243	0.00098	0.00098
1	0.15625	0.18750	0.07680	0.08704	0.02835	0.03078	0.01465	0.01563
2	0.31250	0.50000	0.23040	0.31744	0.13230	0.16308	0.08789	0.10352
3	0.31250	0.81250	0.34560	0.66304	0.30870	0.47178	0.26367	0.36719
4	0.15625	0.96875	0.25920	0.92224	0.36015	0.83193	0.39551	0.76270
5	0.03125	1.00000	0.07776	1.00000	0.16807	1.00000	0.23730	1.00000

k	p = 0.80		p = 5/6		p = 0.90		p = 0.95	
	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$
0	0.00032	0.00032	0.00013	0.00013	0.00001	0.00001	0.00000	0.00000
1	0.00640	0.00672	0.00322	0.00334	0.00045	0.00046	0.00003	0.00003
2	0.05120	0.05792	0.03215	0.03549	0.00810	0.00856	0.00113	0.00116
3	0.20480	0.26272	0.16075	0.19624	0.07290	0.08146	0.02143	0.02259
4	0.40960	0.67232	0.40188	0.59812	0.32805	0.40951	0.20363	0.22622
5	0.32768	1.00000	0.40188	1.00000	0.59049	1.00000	0.77378	1.00000

n = 6

k	p = 0.05		p = 0.10		p = 1/6		p = 0.20	
	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$	$B_{n,p}(\{k\})$	$B_{n,p}(\{0;\ldots;k\})$
0	0.73509	0.73509	0.53144	0.53144	0.33490	0.33490	0.26214	0.26214
1	0.23213	0.96723	0.35429	0.88574	0.40188	0.73678	0.39322	0.65536
2	0.03054	0.99777	0.09842	0.98415	0.20094	0.93771	0.24576	0.90112
3	0.00214	0.99991	0.01458	0.99873	0.05358	0.99130	0.08192	0.98304
4	0.00008	1.00000	0.00122	0.99995	0.00804	0.99934	0.01536	0.99840
5	0.00000	1.00000	0.00005	1.00000	0.00064	0.99998	0.00154	0.99994
6	0.00000	1.00000	0.00000	1.00000	0.00002	1.00000	0.00006	1.00000

Juan J. Cuadrado Gallego

T4. Modelos Univariantes de Distribución

Función de Probabilidad de Geométrica: Número de elementos hasta el primer defectuoso.

$$p(n) = P(x = n) = p(1 - p)^{n-1}; con\ n = 1, \dots\ //\ \mu_x = \frac{q}{p}\ //\ \sigma_x^2 = \frac{q}{p^2}$$

T4. Modelos Univariantes de Distribución

Función de Probabilidad de Poisson: Se aplica al cálculo de probabilidad de aparición de *n* elementos de la población en un intervalo de duración, longitud o cantidad, u otra magnitud fija. El proceso es estable, esto es, el número medio de sucesos por unidad de tiempo, longitud, cantidad, etc., es constante.

$$p(r) = P(x = r) = \frac{\lambda^r}{r!} e^{-\lambda}; con\, r = 0, 1, \ldots \; // \; \mu_x = \lambda \; // \; \sigma_x^2 = \lambda$$

Tiene, al menos, dos funciones asociadas: Weibull y Gompertz.

T4. Modelos Univariantes de Distribución

Juan J. Cuadrado Gallego

Función de Densidad: Función matemática continua que especifica la probabilidad de aparición cada valor del espacio muestral de una variable continua.

$p(x_i) = f(x = x_i)$ Como f es continua → $\int_{-\infty}^{\infty} f(x)dx = 1$

Función de Densidad Normal: Cuando los resultados de un experimento son debidos a un conjunto muy grande de causas independientes, que actúan sumando sus efectos, siendo cada efecto individual de poca importancia respecto al conjunto, es esperable que sigan una distribución normal.

$$p(x_i) = f(x = x_i) = \frac{1}{\sigma_x\sqrt{2\pi}} exp\left(-\frac{1}{2\sigma_x^2}(x_i - \mu_x)^2\right)$$

Se verifica que en el intervalo $\mu \pm 2\sigma$ se encuentra el 95% de los valores y en $\mu \pm 3\sigma$ el 99%

Tiene, al menos, tres funciones derivadas: Chi-cuadrado de Pearson, t de Student y F de Fisher

Se denominan funciones derivadas de la normal porque para la construcción de las correspondientes variables se requiere al menos una distribución normal.

T4. Modelos Univariantes de Distribución

Juan J. Cuadrado Gallego

Función de Densidad Normal: $p(x_i) = f(x = x_i) = \frac{1}{\sigma_x\sqrt{2\pi}} exp\left(-\frac{1}{2\sigma_x^2}(x_i - \mu_x)^2\right)$

Para hacer los cálculos más sencillos se puede hacer el cambio de variable: $z = \frac{x - \mu_x}{\sigma_x}$

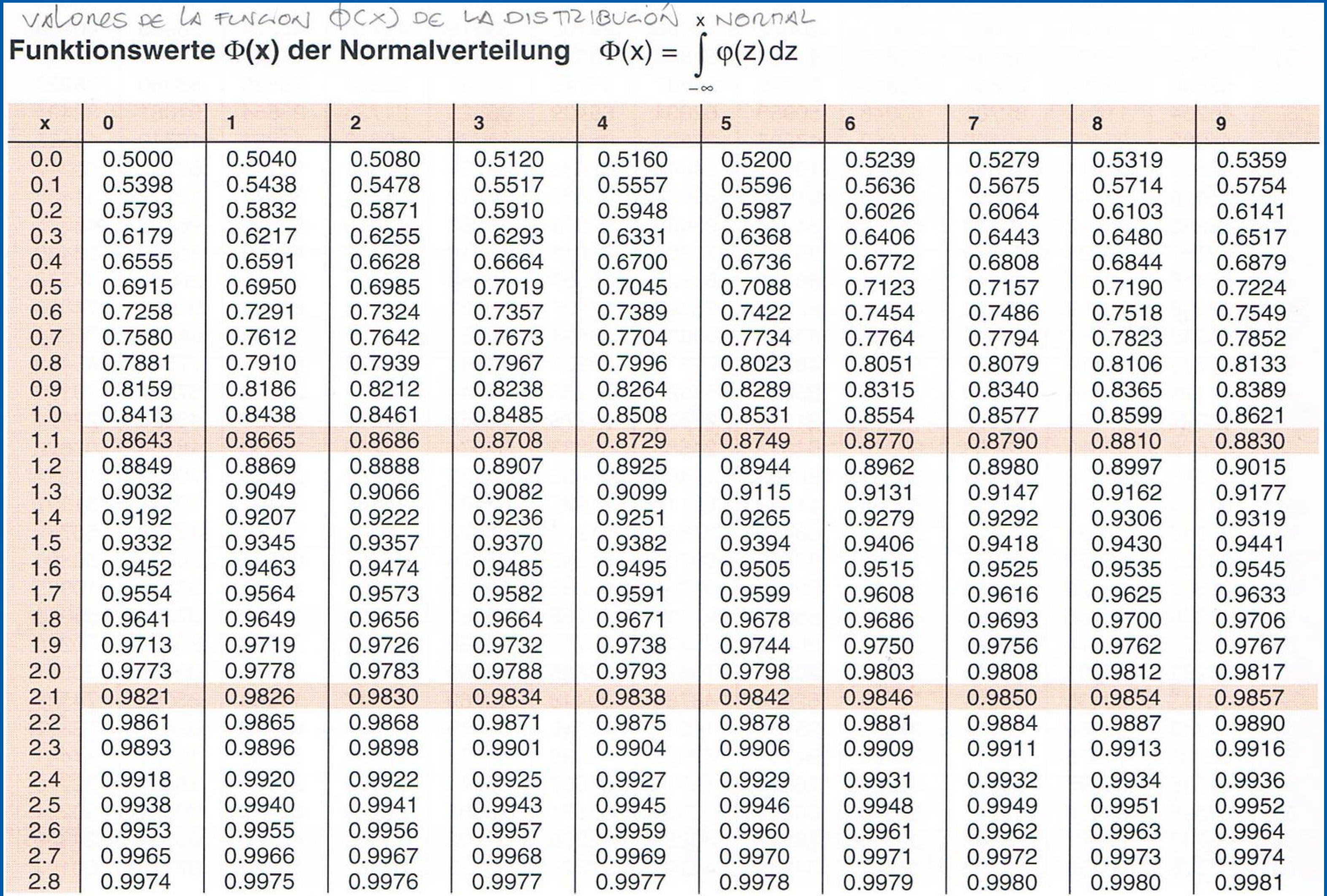
Ya que la variable z sigue una Normal Estándar con $\mu_x = 0$ y $\sigma_x = 1$ con lo que la ecuación de más arriba se convierte en:

$$p(z_i) = f(z = z_i) = \frac{1}{\sigma_x\sqrt{2\pi}} exp\left(-\frac{z_i^2}{2}\right)$$

Y su Función de Distribución está tabulada

T4. Modelos Univariantes de Distribución

La Función de Distribución Normal está tabulada



VALORES DE LA FUNCION Φ(X) DE LA DISTRIBUCIÓN NORMAL

Funktionswerte Φ(x) der Normalverteilung $\Phi(x) = \int_{-\infty}^{x} \varphi(z)\,dz$

x	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5200	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6555	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7045	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981

Juan J. Cuadrado Gallego

T4. Modelos Univariantes de Distribución

Función de Densidad χ^2 (chi o ji cuadrado) de Pearson: Si se tienen n variables aleatorias $x_1 \dots x_n$ independientes y distribuidas N(0, 1), y se define a partir de ellas la variable aleatoria $x = x_1^2 + \cdots + x_n^2$. x sigue una distribución de probabilidad χ^2 de Pearson con n grados de libertad.

La función de densidad es:

$$f(x) = \begin{cases} \dfrac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} si\ x \geq 0 \\ 0\ si\ x < 0 \end{cases}$$

Y su Función de Distribución está tabulada

Juan J. Cuadrado Gallego

T4. Modelos Univariantes de Distribución

Función de Densidad *t* de Student: Si se tienen 2 variables aleatorias independientes, x distribuida N(0, 1) e y distribuida χ_n^2 de Pearson. Si se define la variable $t = \frac{x}{\sqrt{\frac{y}{n}}}$. t sigue una distribución t de Student con n grados de libertad t_n

La función de densidad es:

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} . \frac{(1+\frac{t^2}{n})^{\frac{n+1}{2}}}{\sqrt{n\pi}} \ \forall\ t \in \mathbb{R}$$

Y su Función de Distribución está tabulada

-

T4. Modelos Univariantes de Distribución

Función de Densidad F de Fisher: Si se tienen 2 variables aleatorias independientes, x distribuida χ_m^2 e y distribuida χ_n^2 de Pearson. Si se define la variable $F = \frac{\frac{x}{m}}{\frac{y}{n}} = \frac{xn}{ym}$ sigue una distribución F de Fisher con m grados de libertad en el numerador y n grados de libertad en el denominador $F_{m,n}$

La función de densidad es:

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{F^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}F\right)^{\frac{m+n}{2}}} \, si \; F \geq 0 \\ 0 \; si \; F < 0 \end{cases}$$

Y su Función de Distribución está tabulada

Juan J. Cuadrado Gallego

Universidad de Alcalá

Juan J. Cuadrado Gallego

Estadística-780004

Unidad Temática 3. Inferencia Estadística

Tema 4. Variables Aleatorias

T4. Variables Aleatorias

Estudio de la Población a partir de la Distribución de Probabilidad de las Variables

No se conocen los valores exactos de las variables, como es estadística descriptiva, pero, a través del estudio de una muestra, se han determinado sus probabilidades de aparición.

[1] define variable (aleatoria) como: estadística. 1. f. Mat. Magnitud cuyos valores están determinados por las leyes de probabilidad, como los puntos resultantes de la tirada de un dado.

En todo proceso de observación o experimento se puede definir una variable aleatoria asignando a cada resultado del experimento un número:

- Si el resultado es numérico porque se cuenta o mide el valor
- Si el resultado es cualitativo, porque se le hace corresponder un número arbitrariamente

T4. Variables Aleatorias

Como se vio en Estadística Descriptiva, las variables pueden ser discretas y continuas.

Función de Probabilidad: Función matemática discreta que especifica la probabilidad de aparición cada valor del espacio muestral de una variable discreta.

$$p(x_i) = P(x = x_i) \quad \text{Si E es el espacio muestral } \rightarrow \sum_{i \in E} p(x_i) = 1$$

Función de Densidad: Función matemática continua que especifica la probabilidad de aparición cada valor del espacio muestral de una variable continua.

$$p(x_i) = f(x = x_i) \quad \text{Como f es continua } \rightarrow \int_{-\infty}^{\infty} f(x)dx = 1$$

Función de Distribución: Función matemática discreta o continua que especifica para cada valor del espacio muestral de una variable discreta o continua la probabilidad de aparición un valor menor o igual.

$$F(x_i) = P(x \leq x_i)$$

- Si es una variable discreta: $F(x_k) = P(x \leq x_k) = \sum_{i=1}^{k} p(x_i)$ con n valores $\sum_{i=1}^{k} p(x_i) = 1$
- Si es una variable continua: $F(x_k) = P(x \leq x_k) = \int_{-\infty}^{x_k} f(x)dx$

Juan J. Cuadrado Gallego

T1. Descripción de una variable (Recordatorio)

En el primer análisis de los datos se cuantifica la frecuencia de aparición de los mismos.

[1] define Frecuencia como:

3.f.Estad. Número de elementos comprendidos dentro de un intervalo en una distribución determinada.

En Estadística dicha definición se amplia y profundiza a través de los siguientes conceptos:

Frecuencia Absoluta f_i: Número de apariciones de un dato dado.

Frecuencia Relativa f_{ri}: Número de apariciones de un dato dado dividido entre el número de datos.

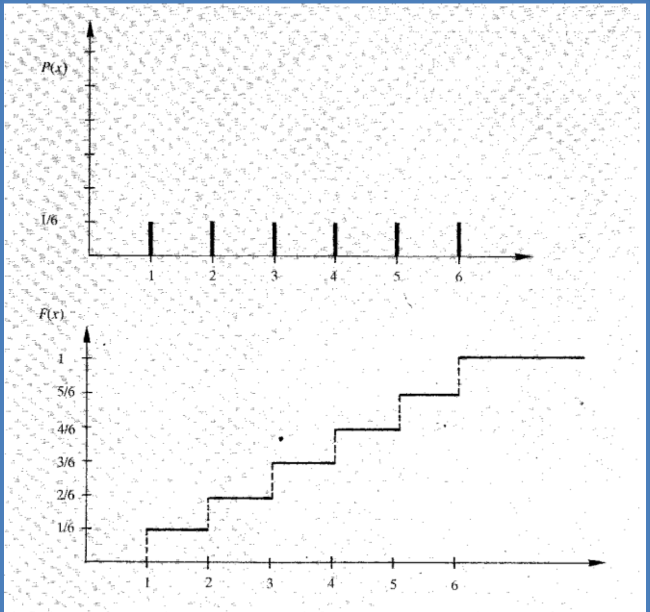
Frecuencia Acumulada (Absoluta o Relativa): Con los datos ordenados por magnitud, suma de las frecuencias absolutas o relativas de los datos inferiores al dato más la del dato.

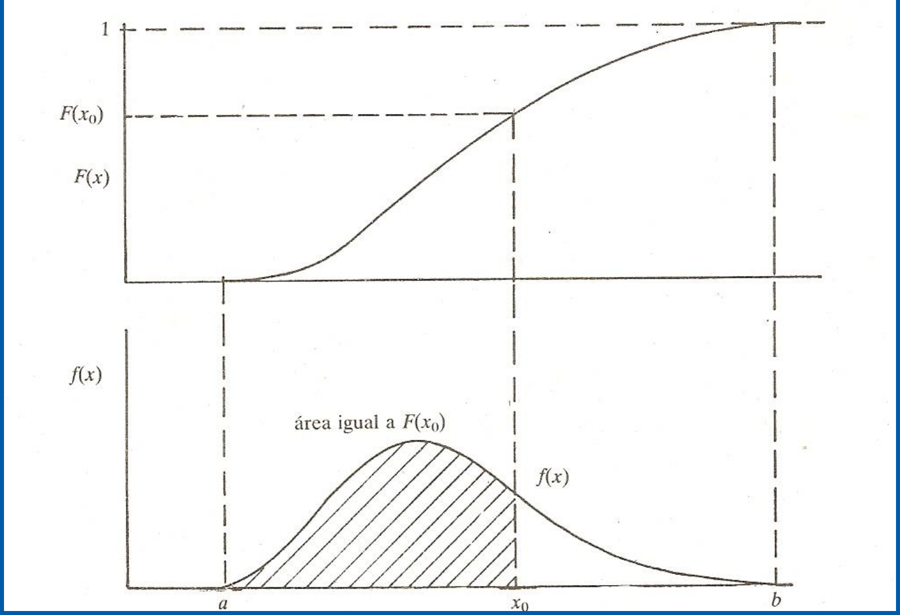
Distribución de Frecuencias (Absoluta o Relativa): Pares compuestos de cada dato y su frecuencias.

Para datos agrupados en clases de equivalencia, el concepto de dato se cambia por el de clase.

T4. Variables Aleatorias

Juan J. Cuadrado Gallego





T4. Variables Aleatorias

Se tiene la siguiente función de densidad de la probabilidad:

$$f(x) = \begin{cases} m.(1 - x^4\ )\ si\ 0 < x < 1 \\ 0\ con\ cualquier\ otro\ valor \end{cases}$$

a) Calcular (1) m y la (2) Función de distribución acumulativa.

a). 1. Solución

$$\int_{-\infty}^{\infty} f(x)dx = 1 \rightarrow \int_{0}^{1} m(1 - x^4)dx = 1 \rightarrow mx|_0^1 - \frac{mx^5}{5}|_0^1 = 1$$

$$(m.1 - m.0) - \left(m.\frac{1^5}{5} - m.\frac{0^5}{5}\right) = 1$$

$$m - \frac{m}{5} = 1 \rightarrow m = \frac{5}{4} = 1.25$$

(2) Pizarra

Juan J. Cuadrado Gallego

T4. Variables Aleatorias

Cuando se tiene una variable Discreta o Continua, cuyos valores se pueden inferir con un grado de probabilidad mediante una Función de Probabilidad o una Función de Densidad, respectivamente, se pueden inferir también los mismos parámetros de caracterización de la población que se obtenían para caracterizar poblaciones en Estadística Descriptiva. Éstos son la media (aquí denominada Esperanza) y la varianza.

Esperanza (o media):

- Variable Discreta: $\mu_x = E(x) = \sum_{i=1}^{n} x_i . p(x_i)$
- Variable Continua: $\mu_x = E(x) = \int_{-\infty}^{\infty} x . f(x) dx$

Varianza:

- Variable Discreta: $\sigma_x^2 = \sum_{i=1}^{\infty} (x_i - \mu)^2 . p(x_i)$
- Variable Continua: $\sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu)^2 . f(x) dx = \int_{-\infty}^{\infty} x^2 . f(x) dx - \mu_x^2$

T4. Variables Aleatorias

Se tiene la siguiente función de densidad de la probabilidad:

$$f(x) = \begin{cases} m.(1 - x^4) \; si \; 0 < x < 1 \\ 0 \; con \; cualquier \; otro \; valor \end{cases}$$

a) Calcular (1) m y la (2) Función de distribución acumulativa.

b) Calcular (1) la media y (2) la varianza de la variable.

$$\mu_x = E(x) = \int_{-\infty}^{\infty} x.f(x)dx$$

$$= \int_0^1 x.1.25.(1 - x^4)dx = 1.25.\left(\frac{x^2}{2} - \frac{x^6}{6}\right)|_0^1 = 0.4167$$

$$\sigma_x^2 = \int_{-\infty}^{\infty} x^2.f(x)dx - \mu_x^2$$

$$= \int_0^1 x^2\left(1.25.(1 - x^4)\right)dx - (0.4167)^2 = 1.25.\left(\frac{x^3}{3} - \frac{x^7}{7}\right)|_0^1 - (0.4167)^2$$

$$= 0.0645$$

Juan J. Cuadrado Gallego

Universidad de Alcalá

Juan J. Cuadrado Gallego

Estadística-780004

Unidad Temática 3. Inferencia Estadística

Tema 5. Estimación puntual y por intervalos

T5. Estimación puntual y por intervalos

Juan J. Cuadrado Gallego

Cuando no se tienen, o no se puede trabajar, con todos los datos de una población, las conclusiones a las que se llegue sobre los mismos solo se podrán dar con una probabilidad.

A dichas conclusiones se llegará mediante el estudio de una o más muestras de datos pertenecientes a la población o/y mediante el conocimiento de la distribución de probabilidad de la variable. Las conclusiones a las que se llegue sobre los datos de cada muestra serán completamente ciertas para dicha muestra; y es cuando dichas conclusiones se trasladan a toda la población cuando aparece la incertidumbre que se cuantifica mediante la probabilidad.

Notación:

- Muestra. Letras latinas $(\bar{x}, s, s^2, r, \ldots)$
- Población. Letras griegas $(\mu, \sigma, \sigma^2, \rho, \ldots)$

T5. Estimación puntual y por intervalos

Hay dos grandes tipos de inferencias que se pueden hacer sobre una población a partir del análisis muestral:

- Estimación. Se trata de inferir el valor de una o más variables descriptoras de la población, por ejemplo la media. Hay dos tipos de estimaciones:
  - Estimación puntual. Se trata de inferir un solo valor para la variable.
  - Estimación por intervalos. Se trata de inferir un intervalo de valores para la variable.
- Contraste de Hipótesis. Se trata de inferir el grado de certeza con que se puede aceptar, o rechazar, una hipótesis formulada sobre una población. Hay dos tipos de contrastes de hipótesis:
  - Paramétrico. Si la población pertenece a una familia conocida de distribuciones.
  - No paramétrico. Si no se sabe a que familia de distribuciones pertenece la población

T5. Estimación puntual y por intervalos

Juan J. Cuadrado Gallego

[1] define muestreo como: estadística. 1. m. Acción de escoger muestras representativas de la calidad o condiciones medias de un todo. 2. m. Técnica empleada para esta selección. 3. m. Selección de una pequeña parte estadísticamente determinada, utilizada para inferir el valor de una o varias características del conjunto.

Un estimador es un parámetro representativo de la población, por ejemplo la media o la varianza.

La media muestral es una variable aleatoria, cuyo valor cambia de muestra en muestra. Si se toman k muestras de tamaño n, cada una de ellas con una media x_i, se tiene un conjunto cuyos elementos son las k medias obtenidas en el muestreo, a partir de las cuales se puede obtener la media de la población. La media poblacional es un valor constante porque solo hay una población.

T5. Estimación puntual y por intervalos

Juan J. Cuadrado Gallego

La clave de la correcta ejecución del muestreo es que las muestras sean representativas de la población.

Métodos de muestreo:

- Muestreo Aleatorio Simple: Cada elemento de la población tiene la misma probabilidad de ser elegido. Las observaciones se realizan con reemplazamiento, de manera que la población es siempre idéntica. Se numeran los elementos de la población de 1 a N y se toman números aleatorios de tantas cifras como tenga N. El valor del número aleatorio indica el valor a seleccionar.

- Muestreo Estratificado: Los elementos de la población son divididos en clases o estratos. Se asigna un número de elementos de la muestra a cada estrato y se escogen los miembros por muestreo aleatorio simple dentro de cada estrato.

- Muestreo por Conglomerados: No se puede aplicar el aleatorio simple ni el estratificado porque no hay una lista con el número de elementos de la población ni de los posibles estratos, pero los elementos se encuentran agrupados en conglomerados que sí se conocen. Se extraen varios conglomerados al azar y se realizan muestreos aleatorios simples dentro de cada uno de ellos.

T5. Estimación puntual y por intervalos

Inferencia o Estimación Puntual para la población a partir de una muestra de tamaño *n*. La población de distribuye $N(0, \sigma^2)$ con varianza conocida.

Media Aritmética:

$$\mu = \bar{x} \qquad \text{Var}(\bar{x}) = \frac{\sigma^2}{n}$$

Varianza:

$$\sigma^2 = s^2$$

T5. Estimación puntual y por intervalos

Juan J. Cuadrado Gallego

Inferencia o Estimación por Intervalos de la media aritmética:

No sabemos que tipo de distribución tiene la población y conocemos la varianza σ^2 poblacional

Partimos de la desigualdad de Tchebychev: $p[\bar{x} - k\sigma_{\bar{x}} < \mu < \bar{x} + k\sigma_{\bar{x}}] \geq 1 - \frac{1}{k^2}$

Sabemos que al ser $(\bar{x})$ una variable aleatoria verifica $\mathrm{E}(\bar{x}) = \mu$ y Var$(\bar{x}) = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ siendo n el número de datos de la muestra.

Sustituyendo tenemos: $p\left[\bar{x} - k\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + k\frac{\sigma}{\sqrt{n}}\right] \geq 1 - \frac{1}{k^2}$

Si se toma el nivel de confianza $1 - \alpha = 1 - \frac{1}{k^2} \rightarrow k = \frac{1}{\sqrt{n}}$

Si se sustituye k, p es: $p\left[\bar{x} - \frac{\sigma}{\sqrt{\alpha n}} < \mu < \bar{x} + \frac{\sigma}{\sqrt{\alpha n}}\right] \geq 1 - \alpha$

Por lo que el intervalo para la media μ con un nivel de confianza $1 - \alpha$ es:

$$\left(\bar{x} - \frac{\sigma}{\sqrt{\alpha n}}, \bar{x} + \frac{\sigma}{\sqrt{\alpha n}}\right)$$

T5. Estimación puntual y por intervalos

Inferencia o Estimación por Intervalos de la media aritmética:

No sabemos que tipo de distribución tiene la población, conocemos la varianza σ^2 poblacional y n es suficientemente grande

La variable: $\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ sigue una distribución normal N(0, 1)

Por lo que, para un nivel de confianza $1-\alpha$ tenemos: $p\left[-\lambda_{\alpha/2} < \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < \lambda_{\alpha/2}\right] = 1-\alpha$

Despejando μ tenemos: $p\left[\bar{x} - \lambda_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + \lambda_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \geq 1-\alpha$

Por lo que el intervalo para la media μ con un nivel de confianza $1-\alpha$ es:

$$\left(\bar{x} - \lambda_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + \lambda_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

n es "suficientemente grande": Es arbitrario pero normalmente es $n > 30$

Se utiliza n "muy grande" para $n > 100$

T5. Estimación puntual y por intervalos

Inferencia o Estimación por Intervalos de la media aritmética:

La población sigue una distribución normal y conocemos la varianza σ^2 poblacional.

El intervalo para la media μ con un nivel de confianza $1 - \alpha$ es, como el anterior:

$$\left(\bar{x} - \lambda_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + \lambda_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

Parece ilógico que si se desconoce la media poblacional se pueda conocer la varianza poblacional pero si suele suceder en los estudios reales, ya que se puede tener información de grandes muestras de poblaciones similares desde las que se puede extrapolar la varianza

T5. Estimación puntual y por intervalos

Inferencia o Estimación por Intervalos de la media aritmética:

La población sigue una distribución normal y no conocemos la varianza σ^2 poblacional.

La variable: $\frac{\bar{x}-\mu}{s}\sqrt{n-1}$ sigue una distribución de probabilidad t de student con n-1 grados de libertad

Por lo que, para un nivel de confianza $1-\alpha$ tenemos: $p\left[-t_{\alpha/2} < \frac{\bar{x}-\mu}{s}\sqrt{n-1} < t_{\alpha/2}\right] = 1-\alpha$

Despejando μ tenemos: $p\left[\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n-1}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n-1}}\right] \geq 1-\alpha$

Por lo que el intervalo para la media μ con un nivel de confianza $1-\alpha$ es:

$$\left(\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n-1}}, \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n-1}}\right)$$

n es suficientemente grande se puede hacer $\sigma = s$ y resolverlo con una distribución normal; y n, en lugar de n-1

T5. Estimación puntual y por intervalos

Inferencia o Estimación por Intervalos de la varianza:

La población sigue una distribución normal y conocemos la media μ poblacional.

Sabemos que: $\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2}$ sigue aproximadamente una distribución de probabilidad chi cuadrado con n grados de libertad χ_n^2

El estimador máximo verosímil de $\boldsymbol{\sigma}^2$ es: $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$

Por lo que la variable $\frac{n\hat{\sigma}^2}{\sigma^2}$ sigue aproximadamente una distribución chi cuadrado con n grados de libertad χ_n^2 y para un nivel de confianza $1-\alpha$ tenemos:

$$p\left[\chi^2_{1-\frac{\alpha}{2}} < \frac{n\hat{\sigma}^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}\right] = 1-\alpha$$

Despejando σ^2 e invirtiendo la desigualdad:

$$p\left[\frac{n\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{n\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2}}}\right] = 1-\alpha$$

Por lo que el intervalo para varianza con un nivel de confianza $1-\alpha$ es:

$$\left(\frac{n\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}}}, \frac{n\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2}}}\right)$$

Juan J. Cuadrado Gallego

T5. Estimación puntual y por intervalos

Inferencia o Estimación por Intervalos de la varianza:

La población sigue una distribución normal y no conocemos la media μ poblacional.

La variable $\frac{ns^2}{\sigma^2}$ sigue aproximadamente una distribución chi cuadrado con n-1 grados de libertad χ^2_{n-1} y para un nivel de confianza $1-\alpha$ tenemos:

$$p\left[\chi^2_{1-\frac{\alpha}{2}} < \frac{ns^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}\right] = 1-\alpha$$

Despejando σ^2 e invirtiendo la desigualdad:

$$p\left[\frac{ns^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{ns^2}{\chi^2_{1-\frac{\alpha}{2}}}\right] = 1-\alpha$$

Por lo que el intervalo para varianza con un nivel de confianza $1-\alpha$ es:

$$\left(\frac{ns^2}{\chi^2_{\frac{\alpha}{2}}}, \frac{ns^2}{\chi^2_{1-\frac{\alpha}{2}}}\right)$$

T5. Estimación puntual y por intervalos

Inferencia o Estimación por Intervalos de la correlación lineal:

Las poblaciones siguen una distribución normal.

La variable: $\dfrac{\frac{1}{2}ln\left(\frac{1+r}{1-r}\right)-\frac{1}{2}ln\left(\frac{1+\rho}{1-\rho}\right)}{\sqrt{\frac{1}{n-3}}}$

para n suficientemente grande sigue aproximadamente una distribución N(0, 1) y para un nivel de confianza $1-\alpha$ tenemos:

$$p\left[\frac{1}{2}ln\left(\frac{1+\rho}{1-\rho}\right) \in \left[\frac{1}{2}ln\left(\frac{1+r}{1-r}\right) -\lambda_{\alpha/2}\sqrt{\frac{1}{n-3}}, \quad \frac{1}{2}ln\left(\frac{1+r}{1-r}\right) +\lambda_{\alpha/2}\sqrt{\frac{1}{n-3}}\right]\right] = 1-\alpha$$

de donde obtenemos el intervalo de confianza para ρ

T5. Estimación puntual y por intervalos

Determinación del tamaño muestral para la estimación de la media de una población normal con varianza conocida.

Antes de obtener una muestra para un estimador es muy útil determinar el tamaño muestral que nos permitirá obtener el intervalo de confianza deseado. Este problema no tiene solución teórica para todos los estimadores, pero sí en este caso.

Sabemos que el intervalo de confianza es: $(\bar{x} - \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}})$

Por lo que el error de estimación es: $e = 2\lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

de donde despejando obtenemos n: $n = 4\left(\lambda_{\alpha/2}\right)^2 \frac{\sigma^2}{e^2}$

Universidad de Alcalá

Estadística-780004

Unidad Temática 3. Inferencia Estadística

Tema 6. Contraste de Hipótesis

Juan J. Cuadrado Gallego

T6. Contraste de Hipótesis

Juan J. Cuadrado Gallego

[1] define Hipótesis como: 1. f. Suposición de algo posible o imposible para sacar de ello una consecuencia. ~ de trabajo. 1. f. hipótesis que se establece provisionalmente como base de una investigación que puede confirmar o negar la validez de aquella.

Tipos de Hipótesis:

Hipótesis Nula (H_0): hipótesis que se contrasta. El nombre de nula (neutra) indica que la hipótesis que se mantendrá a no ser que los datos indiquen su falsedad.

Hipótesis Alternativa (H_1): Si se rechaza la H0, implícitamente se acepta la hipótesis alternativa.

Para aceptar o rechazar una hipótesis se estudia la distribución de probabilidad de un parámetro.

Las hipótesis pueden ser:

Simples: Determinan un único valor para el parámetro. $H_0 = 7$

Compuestas: Determinan un intervalo de valores para el parámetro. $63 \leq H_0 \leq 345$

T6. Contraste de Hipótesis

Para aceptar o rechazar una hipótesis nula se realiza un Contraste.

Tipos de contraste:

- Bilateral: El valor del parámetro de contraste se encuentra dentro de un intervalo cuyos límites vienen dados por la significación elegida.
- Contraste Unilateral: El valor del parámetro de contraste es mayor o menor, dependiendo del caso, que un valor dado por la significación elegida.

El parámetro de contraste será el mismo en ambos tipos y solo dependerá de la variable que se desea contrastar.

T6. Contraste de Hipótesis

Pasos de la metodología de contraste:

1. Se definen H_0 y H_1
2. Se calcula la medida de discrepancia *d* entre los datos muestrales *x* y H_0
3. Se obtienen las discrepancias incompatibles la aceptación de H_0 con nivel de significación α
4. Se analiza *d* y se decide aceptar o no H_0

T6. Contraste de Hipótesis

Tipos de posible Errores en un contraste de hipótesis:

- Error tipo I. Se rechaza H0 siendo cierta
- Error tipo II. Se acepta H0 siendo falsa

T6. Contraste de Hipótesis

Contraste de la media aritmética:

La población sigue una distribución normal y conocemos la varianza σ^2 poblacional

Las Hipótesis son: $H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0$

Hacemos el contraste sobre la variable $d(\bar{x}, \mu) = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

(que sigue una distribución de probabilidad $N(0,1)$)

Si H_0 es cierta: α es la probabilidad de rechazarla y $1-\alpha$ es la probabilidad de aceptarla. Siendo α el nivel de significación del contraste.

Sabemos que: $p\left(-\lambda_{\alpha/2} < \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < \lambda_{\alpha/2}\right) = 1-\alpha$

Por lo que la región de aceptación de H_0 con probabilidad p es:

$-d_c < d_0 < d_c$ siendo $d_c = \lambda_{\alpha/2}$

T6. Contraste de Hipótesis

Juan J. Cuadrado Gallego

Contraste de la media aritmética:

La población sigue una distribución normal, no conocemos la varianza σ^2 poblacional, pero la muestra es suficientemente grande.

σ^2 se aproxima a s^2 y se hace el mismo análisis que en el caso de σ^2 conocida

T6. Contraste de Hipótesis

Contraste de la media aritmética:

La población sigue una distribución normal, no conocemos la varianza σ^2 poblacional, pero la muestra no es suficientemente grande.

Las Hipótesis son: $H_0 : \mu = \mu_0 \quad H_1 : \mu \neq \mu_0$

Hacemos el contraste sobre la variable $d(\bar{x}, \mu) = \frac{\bar{x} - \mu}{s} \sqrt{n-1}$

(sigue una distribución de probabilidad t de student con n-1 grados de libertad)

Si H_0 es cierta: α es la probabilidad de rechazarla y $1 - \alpha$ es la probabilidad de aceptarla. Siendo α el nivel de significación del contraste.

Sabemos que: $p\left(-t_{\alpha/2} < \frac{\bar{x}-\mu}{s}\sqrt{n-1} < t_{\alpha/2}\right) = 1 - \alpha$

Por lo que la región de aceptación de H_0 con probabilidad p es:

$-d_c < d_0 < d_c$ siendo $d_c = t_{\alpha/2}$

T6. Contraste de Hipótesis

Juan J. Cuadrado Gallego

Contraste de la diferencia de medias de dos poblaciones:

Las poblaciones siguen una distribución normal y conocemos las varianzas poblacionales

Las Hipótesis son: $H_0\colon \mu_x - \mu_y = \mu_0 \quad H_1\colon \mu_x - \mu_y \neq \mu_0$

Hacemos el contraste sobre la variable: $d(\bar{x}, \mu_0) = \frac{\bar{x}-\bar{y}-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}}$

(que sigue una distribución de probabilidad $N(0,1)$)

Si H_0 es cierta: α es la probabilidad de rechazarla y $1-\alpha$ es la probabilidad de aceptarla. Siendo α el nivel de significación del contraste.

Sabemos que: $p\left(-\lambda_{\alpha/2} < \frac{\bar{x}-\bar{y}-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}} < \lambda_{\alpha/2}\right) = 1-\alpha$

Por lo que la región de aceptación de H_0 con probabilidad p es:

$-d_c < d_0 < d_c$ siendo $d_c = \lambda_{\alpha/2}$

T6. Contraste de Hipótesis

Juan J. Cuadrado Gallego

Contraste sobre el coeficiente de correlación:

Las poblaciones siguen una distribución normal y ρ es el coeficiente de correlación entre ambas.

Las Hipótesis son: H_0: $\rho = \rho_0$ $H_1: \rho \neq \rho_0$ (con $\rho_0 \neq 0$)

Hacemos el contraste sobre la variable:

$$d(\bar{x}, r) = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$$

(que sigue una distribución de probabilidad N $\left(\frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right); \sqrt{\frac{1}{n-3}}\right)$)

Si H_0 es cierta: α es la probabilidad de rechazarla y $1-\alpha$ es la probabilidad de aceptarla. Siendo α el nivel de significación del contraste.

Sabemos que: $p\left(\frac{1}{2}\ln\left(\frac{1+\rho_0}{1-\rho_0}\right) - \lambda_{\frac{\alpha}{2}}\sqrt{\frac{1}{n-3}} < d_0 < \frac{1}{2}\ln\left(\frac{1+\rho_0}{1-\rho_0}\right) + \lambda_{\frac{\alpha}{2}}\sqrt{\frac{1}{n-3}}\right) = 1-\alpha$

Por lo que la región de aceptación de H_0 con probabilidad p es:

$$-d_c < d_0 < d_c \text{ siendo } d_c = \frac{1}{2}\ln\left(\frac{1+\rho_0}{1-\rho_0}\right) - \lambda_{\frac{\alpha}{2}}\sqrt{\frac{1}{n-3}}$$

T6. Contraste de Hipótesis

Contraste sobre la incorrelación: Las poblaciones siguen una distribución normal y ρ es el coeficiente de correlación entre ambas.

Las Hipótesis son: H_0: $\rho = 0$ H_1: $\rho \neq 0$

Hacemos el contraste sobre las variables:

- $\frac{r^2}{1-r^2}(n-2)$ (que sigue una distribución F ($1; n-2$) (numerador, denominador)
- $\sqrt{\frac{r^2}{1-r^2}(n-2)}$ (que sigue una distribución t ($n-2$)

Si H_0 es cierta: α, que es el nivel de significación, es la probabilidad de rechazarla y $1-\alpha$ es la probabilidad de aceptarla. Sabemos que:

- $p\left[0 < \frac{r^2}{1-r^2}(n-2) < F_{(1,n-2),1-\alpha}\right[= 1-\alpha$
- $p\left(-t_{\frac{\alpha}{2}} < \sqrt{\frac{r^2}{1-r^2}(n-2)} < t_{\frac{\alpha}{2}}\right) = 1-\alpha$

Por lo que la región de aceptación de H_0 con probabilidad p es:

Unilateral: $0 < d_0 < d_c$ siendo $d_c = F_{(1,n-2),1-\alpha}$

Bilateral: $-d_c < d_0 < d_c$ siendo $d_c = t_{\frac{\alpha}{2}}$

T6. Contraste de Hipótesis

Contraste de la proporción de una característica:

Las Hipótesis son: $H_0 : p = p \quad H_1 : p \neq p_0$

Hacemos el contraste sobre la variable:

$$d(\bar{x}, \hat{p}) = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$ (que sigue una distribución N $(0{,}1)$

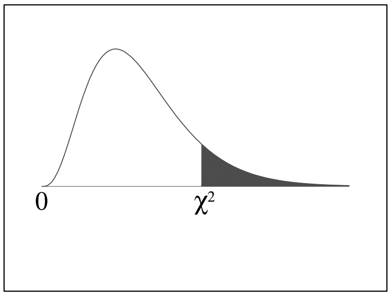
Si H_0 es cierta: α es la probabilidad de rechazarla y $1 - \alpha$ es la probabilidad de aceptarla. Sabemos que:

Sabemos que: $p\left(-\lambda_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} < \lambda_{\alpha/2}\right) = 1 - \alpha$

Por lo que la región de aceptación de H_0 con probabilidad *p* es:

$-d_c < d_0 < d_c$ siendo $d_c = \lambda_{\alpha/2}$

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Table entry for *p* is the critical value *F** with probability *p* lying to its right.

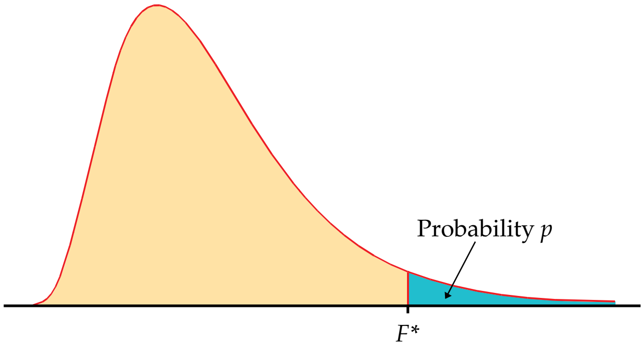


TABLE E

F critical values

			Degrees of freedom in the numerator								
		p	1	2	3	4	5	6	7	8	9
Degrees of freedom in the denominator	1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
		.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
		.025	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
		.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
		.001	405284	500000	540379	562500	576405	585937	592873	598144	602284
	2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
		.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
		.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
		.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
		.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
	3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
		.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
		.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
		.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
		.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
	4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
		.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
		.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
		.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
		.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
	5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
		.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
		.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
		.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
		.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
	6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
		.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
		.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
		.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
		.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
	7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
		.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
		.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
		.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
		.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33

Table entry for p is the critical value F^* with probability p lying to its right.

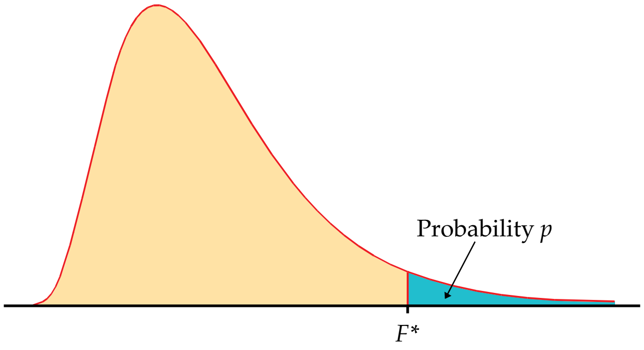


TABLE E

F critical values (continued)

Degrees of freedom in the numerator										
10	12	15	20	25	30	40	50	60	120	1000
60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.69	62.79	63.06	63.30
241.88	243.91	245.95	248.01	249.26	250.10	251.14	251.77	252.20	253.25	254.19
968.63	976.71	984.87	993.10	998.08	1001.4	1005.6	1008.1	1009.8	1014.0	1017.7
6055.8	6106.3	6157.3	6208.7	6239.8	6260.6	6286.8	6302.5	6313.0	6339.4	6362.7
605621	610668	615764	620908	624017	626099	628712	630285	631337	633972	636301
9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.47	9.48	9.49
19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48	19.48	19.49	19.49
39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.48	39.49	39.50
99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.48	99.49	99.50
999.40	999.42	999.43	999.45	999.46	999.47	999.47	999.48	999.48	999.49	999.50
5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15	5.15	5.14	5.13
8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.58	8.57	8.55	8.53
14.42	14.34	14.25	14.17	14.12	14.08	14.04	14.01	13.99	13.95	13.91
27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.35	26.32	26.22	26.14
129.25	128.32	127.37	126.42	125.84	125.45	124.96	124.66	124.47	123.97	123.53
3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.80	3.79	3.78	3.76
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.70	5.69	5.66	5.63
8.84	8.75	8.66	8.56	8.50	8.46	8.41	8.38	8.36	8.31	8.26
14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.69	13.65	13.56	13.47
48.05	47.41	46.76	46.10	45.70	45.43	45.09	44.88	44.75	44.40	44.09
3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.15	3.14	3.12	3.11
4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.44	4.43	4.40	4.37
6.62	6.52	6.43	6.33	6.27	6.23	6.18	6.14	6.12	6.07	6.02
10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.24	9.20	9.11	9.03
26.92	26.42	25.91	25.39	25.08	24.87	24.60	24.44	24.33	24.06	23.82
2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.77	2.76	2.74	2.72
4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.75	3.74	3.70	3.67
5.46	5.37	5.27	5.17	5.11	5.07	5.01	4.98	4.96	4.90	4.86
7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.09	7.06	6.97	6.89
18.41	17.99	17.56	17.12	16.85	16.67	16.44	16.31	16.21	15.98	15.77
2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.52	2.51	2.49	2.47
3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.32	3.30	3.27	3.23
4.76	4.67	4.57	4.47	4.40	4.36	4.31	4.28	4.25	4.20	4.15
6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.86	5.82	5.74	5.66
14.08	13.71	13.32	12.93	12.69	12.53	12.33	12.20	12.12	11.91	11.72

(Continued)

TABLE E

F critical values (continued)

Degrees of freedom in the denominator		Degrees of freedom in the numerator								
	p	1	2	3	4	5	6	7	8	9
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.001	25.41	18.49	15.83	14.39	13.48	12.86	12.40	12.05	11.77
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.001	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.001	21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.001	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.001	17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.001	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.001	16.12	10.97	9.01	7.94	7.27	6.80	6.46	6.19	5.98
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68
	.001	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75

TABLE E

F critical values (continued)

Degrees of freedom in the numerator										
10	12	15	20	25	30	40	50	60	120	1000
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.35	2.34	2.32	2.30
3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.02	3.01	2.97	2.93
4.30	4.20	4.10	4.00	3.94	3.89	3.84	3.81	3.78	3.73	3.68
5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.07	5.03	4.95	4.87
11.54	11.19	10.84	10.48	10.26	10.11	9.92	9.80	9.73	9.53	9.36
2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.22	2.21	2.18	2.16
3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.80	2.79	2.75	2.71
3.96	3.87	3.77	3.67	3.60	3.56	3.51	3.47	3.45	3.39	3.34
5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.52	4.48	4.40	4.32
9.89	9.57	9.24	8.90	8.69	8.55	8.37	8.26	8.19	8.00	7.84
2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.12	2.11	2.08	2.06
2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.64	2.62	2.58	2.54
3.72	3.62	3.52	3.42	3.35	3.31	3.26	3.22	3.20	3.14	3.09
4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.12	4.08	4.00	3.92
8.75	8.45	8.13	7.80	7.60	7.47	7.30	7.19	7.12	6.94	6.78
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.04	2.03	2.00	1.98
2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.51	2.49	2.45	2.41
3.53	3.43	3.33	3.23	3.16	3.12	3.06	3.03	3.00	2.94	2.89
4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.81	3.78	3.69	3.61
7.92	7.63	7.32	7.01	6.81	6.68	6.52	6.42	6.35	6.18	6.02
2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.97	1.96	1.93	1.91
2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.40	2.38	2.34	2.30
3.37	3.28	3.18	3.07	3.01	2.96	2.91	2.87	2.85	2.79	2.73
4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.57	3.54	3.45	3.37
7.29	7.00	6.71	6.40	6.22	6.09	5.93	5.83	5.76	5.59	5.44
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.92	1.90	1.88	1.85
2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.31	2.30	2.25	2.21
3.25	3.15	3.05	2.95	2.88	2.84	2.78	2.74	2.72	2.66	2.60
4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.38	3.34	3.25	3.18
6.80	6.52	6.23	5.93	5.75	5.63	5.47	5.37	5.30	5.14	4.99
2.10	2.05	2.01	1.96	1.93	1.91	1.89	1.87	1.86	1.83	1.80
2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.24	2.22	2.18	2.14
3.15	3.05	2.95	2.84	2.78	2.73	2.67	2.64	2.61	2.55	2.50
3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.22	3.18	3.09	3.02
6.40	6.13	5.85	5.56	5.38	5.25	5.10	5.00	4.94	4.77	4.62
2.06	2.02	1.97	1.92	1.89	1.87	1.85	1.83	1.82	1.79	1.76
2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.18	2.16	2.11	2.07
3.06	2.96	2.86	2.76	2.69	2.64	2.59	2.55	2.52	2.46	2.40
3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.08	3.05	2.96	2.88
6.08	5.81	5.54	5.25	5.07	4.95	4.80	4.70	4.64	4.47	4.33
2.03	1.99	1.94	1.89	1.86	1.84	1.81	1.79	1.78	1.75	1.72
2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.12	2.11	2.06	2.02
2.99	2.89	2.79	2.68	2.61	2.57	2.51	2.47	2.45	2.38	2.32
3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.97	2.93	2.84	2.76
5.81	5.55	5.27	4.99	4.82	4.70	4.54	4.45	4.39	4.23	4.08
2.00	1.96	1.91	1.86	1.83	1.81	1.78	1.76	1.75	1.72	1.69
2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.08	2.06	2.01	1.97
2.92	2.82	2.72	2.62	2.55	2.50	2.44	2.41	2.38	2.32	2.26
3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.87	2.83	2.75	2.66
5.58	5.32	5.05	4.78	4.60	4.48	4.33	4.24	4.18	4.02	3.87

(*Continued*)

TABLE E

F critical values (continued)

Degrees of freedom in the denominator		Degrees of freedom in the numerator								
	p	1	2	3	4	5	6	7	8	9
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.001	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.001	15.08	10.16	8.28	7.27	6.62	6.18	5.85	5.59	5.39
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.001	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24
21	.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.001	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.001	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.001	14.20	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.89
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.001	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.001	13.88	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.71
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.001	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.001	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57

TABLE E

F critical values (continued)

Degrees of freedom in the numerator										
10	12	15	20	25	30	40	50	60	120	1000
1.98	1.93	1.89	1.84	1.80	1.78	1.75	1.74	1.72	1.69	1.66
2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.04	2.02	1.97	1.92
2.87	2.77	2.67	2.56	2.49	2.44	2.38	2.35	2.32	2.26	2.20
3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.78	2.75	2.66	2.58
5.39	5.13	4.87	4.59	4.42	4.30	4.15	4.06	4.00	3.84	3.69
1.96	1.91	1.86	1.81	1.78	1.76	1.73	1.71	1.70	1.67	1.64
2.38	2.31	2.23	2.16	2.11	2.07	2.03	2.00	1.98	1.93	1.88
2.82	2.72	2.62	2.51	2.44	2.39	2.33	2.30	2.27	2.20	2.14
3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.71	2.67	2.58	2.50
5.22	4.97	4.70	4.43	4.26	4.14	3.99	3.90	3.84	3.68	3.53
1.94	1.89	1.84	1.79	1.76	1.74	1.71	1.69	1.68	1.64	1.61
2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.97	1.95	1.90	1.85
2.77	2.68	2.57	2.46	2.40	2.35	2.29	2.25	2.22	2.16	2.09
3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.64	2.61	2.52	2.43
5.08	4.82	4.56	4.29	4.12	4.00	3.86	3.77	3.70	3.54	3.40
1.92	1.87	1.83	1.78	1.74	1.72	1.69	1.67	1.66	1.62	1.59
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.94	1.92	1.87	1.82
2.73	2.64	2.53	2.42	2.36	2.31	2.25	2.21	2.18	2.11	2.05
3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.58	2.55	2.46	2.37
4.95	4.70	4.44	4.17	4.00	3.88	3.74	3.64	3.58	3.42	3.28
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.60	1.57
2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.91	1.89	1.84	1.79
2.70	2.60	2.50	2.39	2.32	2.27	2.21	2.17	2.14	2.08	2.01
3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.53	2.50	2.40	2.32
4.83	4.58	4.33	4.06	3.89	3.78	3.63	3.54	3.48	3.32	3.17
1.89	1.84	1.80	1.74	1.71	1.69	1.66	1.64	1.62	1.59	1.55
2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.88	1.86	1.81	1.76
2.67	2.57	2.47	2.36	2.29	2.24	2.18	2.14	2.11	2.04	1.98
3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.48	2.45	2.35	2.27
4.73	4.48	4.23	3.96	3.79	3.68	3.53	3.44	3.38	3.22	3.08
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.57	1.54
2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.86	1.84	1.79	1.74
2.64	2.54	2.44	2.33	2.26	2.21	2.15	2.11	2.08	2.01	1.94
3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.44	2.40	2.31	2.22
4.64	4.39	4.14	3.87	3.71	3.59	3.45	3.36	3.29	3.14	2.99
1.87	1.82	1.77	1.72	1.68	1.66	1.63	1.61	1.59	1.56	1.52
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.84	1.82	1.77	1.72
2.61	2.51	2.41	2.30	2.23	2.18	2.12	2.08	2.05	1.98	1.91
3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.40	2.36	2.27	2.18
4.56	4.31	4.06	3.79	3.63	3.52	3.37	3.28	3.22	3.06	2.91
1.86	1.81	1.76	1.71	1.67	1.65	1.61	1.59	1.58	1.54	1.51
2.22	2.15	2.07	1.99	1.94	1.90	1.85	1.82	1.80	1.75	1.70
2.59	2.49	2.39	2.28	2.21	2.16	2.09	2.05	2.03	1.95	1.89
3.09	2.96	2.81	2.66	2.57	2.50	2.42	2.36	2.33	2.23	2.14
4.48	4.24	3.99	3.72	3.56	3.44	3.30	3.21	3.15	2.99	2.84
1.85	1.80	1.75	1.70	1.66	1.64	1.60	1.58	1.57	1.53	1.50
2.20	2.13	2.06	1.97	1.92	1.88	1.84	1.81	1.79	1.73	1.68
2.57	2.47	2.36	2.25	2.18	2.13	2.07	2.03	2.00	1.93	1.86
3.06	2.93	2.78	2.63	2.54	2.47	2.38	2.33	2.29	2.20	2.11
4.41	4.17	3.92	3.66	3.49	3.38	3.23	3.14	3.08	2.92	2.78

(Continued)

TABLE E

F critical values (continued)

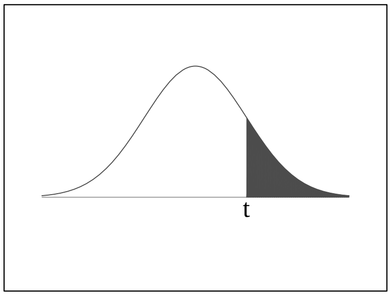
		Degrees of freedom in the numerator								
Degrees of freedom in the denominator	*p*	1	2	3	4	5	6	7	8	9
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.001	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.001	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.001	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.001	12.61	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02
50	.100	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76
	.050	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	.025	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38
	.010	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78
	.001	12.22	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.001	11.97	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
100	.100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69
	.050	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97
	.025	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24
	.010	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
	.001	11.50	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.44
200	.100	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66
	.050	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93
	.025	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18
	.010	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50
	.001	11.15	7.15	5.63	4.81	4.29	3.92	3.65	3.43	3.26
1000	.100	2.71	2.31	2.09	1.95	1.85	1.78	1.72	1.68	1.64
	.050	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89
	.025	5.04	3.70	3.13	2.80	2.58	2.42	2.30	2.20	2.13
	.010	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43
	.001	10.89	6.96	5.46	4.65	4.14	3.78	3.51	3.30	3.13

TABLE E

F critical values (continued)

Degrees of freedom in the numerator										
10	12	15	20	25	30	40	50	60	120	1000
1.84	1.79	1.74	1.69	1.65	1.63	1.59	1.57	1.56	1.52	1.48
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.79	1.77	1.71	1.66
2.55	2.45	2.34	2.23	2.16	2.11	2.05	2.01	1.98	1.91	1.84
3.03	2.90	2.75	2.60	2.51	2.44	2.35	2.30	2.26	2.17	2.08
4.35	4.11	3.86	3.60	3.43	3.32	3.18	3.09	3.02	2.86	2.72
1.83	1.78	1.73	1.68	1.64	1.62	1.58	1.56	1.55	1.51	1.47
2.18	2.10	2.03	1.94	1.89	1.85	1.81	1.77	1.75	1.70	1.65
2.53	2.43	2.32	2.21	2.14	2.09	2.03	1.99	1.96	1.89	1.82
3.00	2.87	2.73	2.57	2.48	2.41	2.33	2.27	2.23	2.14	2.05
4.29	4.05	3.80	3.54	3.38	3.27	3.12	3.03	2.97	2.81	2.66
1.82	1.77	1.72	1.67	1.63	1.61	1.57	1.55	1.54	1.50	1.46
2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.76	1.74	1.68	1.63
2.51	2.41	2.31	2.20	2.12	2.07	2.01	1.97	1.94	1.87	1.80
2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.25	2.21	2.11	2.02
4.24	4.00	3.75	3.49	3.33	3.22	3.07	2.98	2.92	2.76	2.61
1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.42	1.38
2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.66	1.64	1.58	1.52
2.39	2.29	2.18	2.07	1.99	1.94	1.88	1.83	1.80	1.72	1.65
2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.06	2.02	1.92	1.82
3.87	3.64	3.40	3.14	2.98	2.87	2.73	2.64	2.57	2.41	2.25
1.73	1.68	1.63	1.57	1.53	1.50	1.46	1.44	1.42	1.38	1.33
2.03	1.95	1.87	1.78	1.73	1.69	1.63	1.60	1.58	1.51	1.45
2.32	2.22	2.11	1.99	1.92	1.87	1.80	1.75	1.72	1.64	1.56
2.70	2.56	2.42	2.27	2.17	2.10	2.01	1.95	1.91	1.80	1.70
3.67	3.44	3.20	2.95	2.79	2.68	2.53	2.44	2.38	2.21	2.05
1.71	1.66	1.60	1.54	1.50	1.48	1.44	1.41	1.40	1.35	1.30
1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.56	1.53	1.47	1.40
2.27	2.17	2.06	1.94	1.87	1.82	1.74	1.70	1.67	1.58	1.49
2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.88	1.84	1.73	1.62
3.54	3.32	3.08	2.83	2.67	2.55	2.41	2.32	2.25	2.08	1.92
1.66	1.61	1.56	1.49	1.45	1.42	1.38	1.35	1.34	1.28	1.22
1.93	1.85	1.77	1.68	1.62	1.57	1.52	1.48	1.45	1.38	1.30
2.18	2.08	1.97	1.85	1.77	1.71	1.64	1.59	1.56	1.46	1.36
2.50	2.37	2.22	2.07	1.97	1.89	1.80	1.74	1.69	1.57	1.45
3.30	3.07	2.84	2.59	2.43	2.32	2.17	2.08	2.01	1.83	1.64
1.63	1.58	1.52	1.46	1.41	1.38	1.34	1.31	1.29	1.23	1.16
1.88	1.80	1.72	1.62	1.56	1.52	1.46	1.41	1.39	1.30	1.21
2.11	2.01	1.90	1.78	1.70	1.64	1.56	1.51	1.47	1.37	1.25
2.41	2.27	2.13	1.97	1.87	1.79	1.69	1.63	1.58	1.45	1.30
3.12	2.90	2.67	2.42	2.26	2.15	2.00	1.90	1.83	1.64	1.43
1.61	1.55	1.49	1.43	1.38	1.35	1.30	1.27	1.25	1.18	1.08
1.84	1.76	1.68	1.58	1.52	1.47	1.41	1.36	1.33	1.24	1.11
2.06	1.96	1.85	1.72	1.64	1.58	1.50	1.45	1.41	1.29	1.13
2.34	2.20	2.06	1.90	1.79	1.72	1.61	1.54	1.50	1.35	1.16
2.99	2.77	2.54	2.30	2.14	2.02	1.87	1.77	1.69	1.49	1.22

t-Distribution Table

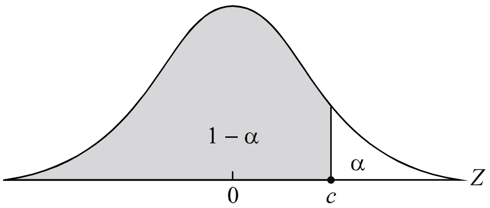


The shaded area is equal to α for $t = t_\alpha$.

df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
32	1.309	1.694	2.037	2.449	2.738
34	1.307	1.691	2.032	2.441	2.728
36	1.306	1.688	2.028	2.434	2.719
38	1.304	1.686	2.024	2.429	2.712
∞	1.282	1.645	1.960	2.326	2.576

TABLA DE LA DISTRIBUCION NORMAL ESTANDAR

La tabla da áreas $1-\alpha$ y valores
$c = Z_{1-\alpha}$, donde, $P[Z \le c] = 1-\alpha$,
y donde Z tiene distribución normal $N(0,1)$.



Z	Segundo decimal de Z									
	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9762	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.8	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.000	1.000	1.000
4.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000