



Estimación de las principales distribuciones de probabilidad mediante Microsoft Excel

Apellidos, nombre	Martínez Gómez, Mónica (momargo@eio.upv.es) Marí Benlloch, Manuel (mamaben@eio.upv.es)
Departamento	Estadística, Investigación Operativa Aplicadas y Calidad
Centro	Universidad Politécnica de Valencia

1. Resumen de las ideas clave

En este artículo vamos a analizar el funcionamiento y aplicaciones de la hoja de cálculo Microsoft Excel, para la estimación de probabilidades de las principales distribuciones de probabilidad, tanto discretas (distribución Binomial y Poisson) como continuas (distribución Normal). La finalidad de ello es realizar una especie de catálogo al que acudir en el caso de que las empresas no dispongan de software estadísticos más específicos.

2. Introducción

Haciendo uso de Microsoft Excel se podrán calcular funciones de distribución a partir de la función de cuantía de una variable aleatoria discreta. También se podrá utilizar Excel para representar las funciones de densidad y distribución de variables continuas en algunos puntos, de manera que se puedan abordar situaciones reales en empresas que carezcan de software estadísticos específicos.

La estructura de este objeto de aprendizaje es como sigue: en primer lugar se presentan los objetivos que pretendemos conseguir; a continuación conoceremos la estimación de probabilidades de los más importantes modelos de probabilidad, definiendo los argumentos de Microsoft Excel y relacionándolos con los conceptos estadísticos correspondientes. Asimismo, comentaremos sus principales características, estableceremos los pasos necesarios para el cálculo y resolveremos algunos ejemplos prácticos para ayudar a su comprensión. Finalmente, destacaremos los conceptos básicos de aprendizaje con respecto al cálculo de probabilidades mediante Microsoft Excel.

3. Objetivos

- Comprender los conceptos teóricos estadísticos.
- Clarificar conceptos respecto al cálculo de probabilidades, de las principales distribuciones: la distribución Binomial y Poisson como discretas y la distribución Normal como continuas.
- Comprobar el procedimiento antes de dar por bueno un resultado.

4. Definición y características de las distribuciones en Microsoft Excel

Actualmente es indiscutible la necesidad de trabajar con programas informáticos en materias relacionadas con la estadística, debido a que muchas veces requieren cálculos tediosos y excesivamente largos. En este sentido, es necesaria la utilización de un determinado software, en nuestro caso Microsoft Excel.

La elección de la hoja de cálculo como herramienta informática frente a programas estadísticos específicos (SPSS, Statgraphics, etc.) estriba sobre todo en su disponibilidad, ya que ésta se puede encontrar en casi todos los ordenadores a los que habitualmente solemos tener acceso y posteriormente en todas las empresas a las que podamos acceder como futuros profesionales, así como en la mayoría de puestos de trabajo. Además, presenta la ventaja de que es mucho más versátil que el software puramente estadístico.

4.1. Distribuciones Discretas

4.1.1. Distribución Binomial

Microsoft Excel nos permite calcular la probabilidad de obtener un número concreto de éxitos ($P(X=x_i)$) de una variable aleatoria discreta que sigue una distribución binomial $B(n,p)$, en problemas con un número fijo de pruebas o ensayos (n), cuando los resultados de un ensayo son sólo éxito o fracaso, cuando los ensayos son independientes y cuando la probabilidad de éxito es constante durante todo el experimento (p).

Para ello, en insertar función, selecciona dentro del tipo de funciones estadísticas, el comando “DISTR.BINOM”, como puede apreciarse en la imagen 1.

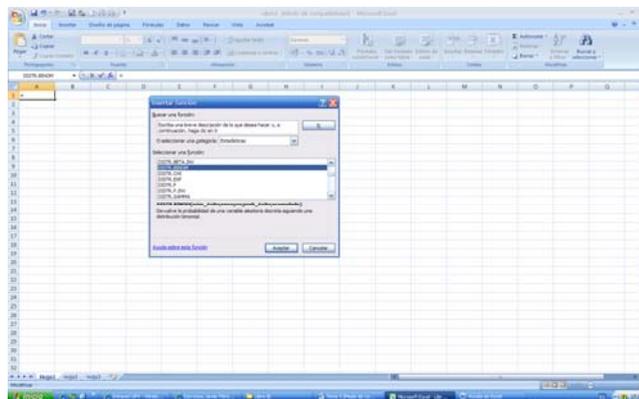


Imagen 1. Pantalla de Microsoft Excel cuando se selecciona la función

Fuente: Elaboración propia

Los argumentos de la función:

- Núm_éxito, es el número de éxitos en los ensayos, que se desean estimar (x_i).
- Ensayos, es el número de ensayos independientes o repeticiones que se realizan (n_i).
- Prob_éxito, es la probabilidad de éxito en cada ensayo (p).
- Acumulado, es un valor lógico que determina la forma de la función. Si el argumento acumulado es VERDADERO, DISTR.BINOM devuelve la función de distribución acumulada, que es la probabilidad de que exista el máximo número de éxitos; si es FALSO, devuelve la función

de masa de probabilidad, que es la probabilidad de que un evento se reproduzca un número de veces igual al argumento núm_éxito.

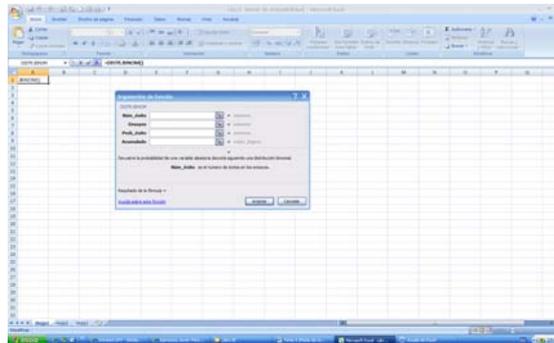


Imagen 2. Pantalla de Microsoft Excel con los argumentos de la Distribución Binomial
Fuente: Elaboración propia

• **Ejemplo:**

En un proceso de fabricación de tornillos se sabe que el 2% son defectuosos. Los empaquetamos en cajas de 50 tornillos. Calcula la probabilidad de que en una caja no haya ningún tornillo defectuoso”.

Pasos:

- En primer lugar determinamos el número de éxitos de los ensayos: 0
- En segundo lugar, definimos el número de ensayos independientes: 50.
- A continuación definimos la probabilidad de éxito de cada ensayo: 0,2.
- Finalmente como nos piden una función de probabilidad o de cuantía, definimos el valor Acumulado como FALSO.

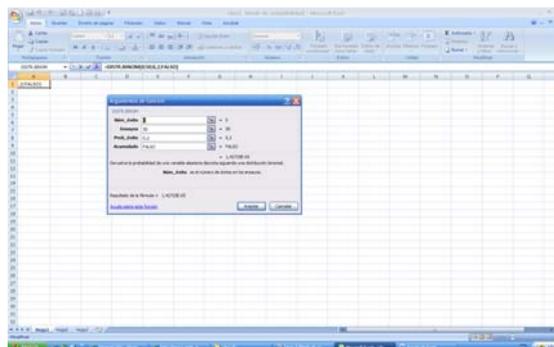


Imagen 3. Pantalla de Microsoft Excel con los argumentos de la Distribución Binomial
Fuente: Elaboración propia

La probabilidad obtenida sería, como se aprecia en la imagen anterior, de 1,42725E-05.

4.1.2. Distribución Poisson

Una de las aplicaciones comunes de la distribución de Poisson es la predicción del número de eventos en un determinado período de tiempo, como por ejemplo, el número de automóviles que se presenta a una zona de peaje en el intervalo de un minuto. Microsoft Office Excel permite calcular dichas probabilidades, introduciendo la función, **fx: POISSON**.

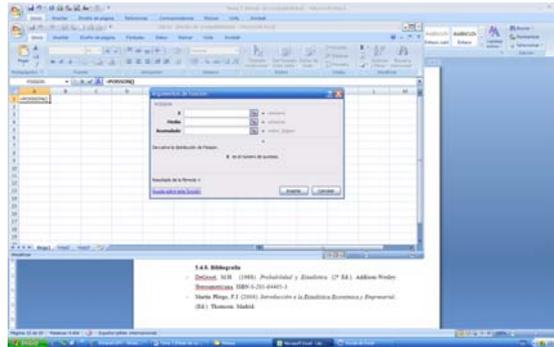


Imagen 4. Pantalla de Microsoft Excel cuando se selecciona la función

Fuente: Elaboración propia

En ella,

- X, es el número de eventos o sucesos (x_i).
- Media, es el valor numérico esperado (valor promedio o λ).
- Acumulado, es un valor lógico que determina la forma de la distribución de probabilidad devuelta. Si el argumento acumulado es VERDADERO, POISSON devuelve la probabilidad de Poisson de que un evento aleatorio ocurra un número de veces comprendido entre 0 y x inclusive; si el argumento acumulado es FALSO, la función devuelve la probabilidad de Poisson de que un evento ocurra exactamente x veces.

- **Ejemplo:**

El número medio de defectos en un rollo de tela es de 0,4. Se inspecciona una muestra de 10 rollos, ¿cuál es la probabilidad de que el número total de defectos en los 10 rollos sea por lo menos de 9?

Pasos:

- En primer lugar definimos la variable ,X ~ Número de defectos por rollo, sigue una distribución de Ps ($\lambda=0,4$).
- Nos solicitan obtener la $P(X \geq 9)$, lo que equivale a estimar, $1- P(X \leq 8)$.

$$P(X \geq 9) = 1 - P(X \leq 8)$$

- Mediante la hoja de cálculo podemos obtener la probabilidad de $P(X \leq 8)$ y sustituir.
- Finalmente como nos piden una función de probabilidad de distribución (probabilidad acumulada), definimos el valor Acumulado como VERDADERO.

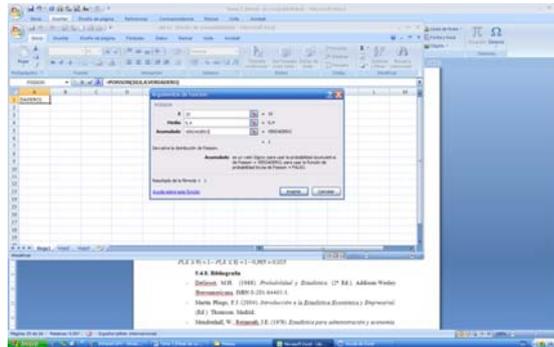


Imagen 5. Pantalla de Microsoft Excel cuando se selecciona la función
Fuente: Elaboración propia

Luego, $P(X \geq 9) = 1 - P(X \leq 8) \approx 1 - 1 = 0$

4.2. Distribuciones Continuas: La Distribución Normal

Microsoft Excel nos permite calcular la probabilidad de una variable aleatoria continua que sigue una distribución normal tanto con una media y desviación típica de cualquier valor, $N(\mu, \sigma)$, como los valores de una normal tipificada o estandarizada, $N(0,1)$. A su vez, también nos permite conocer como se estandarizan o tipifican las variables. Veamos cada caso por separado:

4.2.1.1. Variable Normal $N(\mu, \sigma)$

La función **DISTR.NORM**, permite buscar la probabilidad que en una distribución normal de parámetros $N(\mu, \sigma)$ deja por debajo el valor "a", ($P(X \leq a)$). Para ello, en insertar función, selecciona dentro del tipo de funciones estadísticas, el comando "DISTR.NORM".

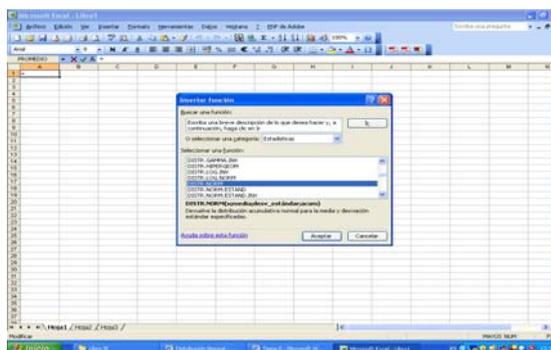


Imagen 6. Pantalla de Microsoft Excel cuando se selecciona la función
Fuente: Elaboración propia

Esta función, devuelve la distribución normal para la media y desviación estándar especificadas. Tiene un gran número de aplicaciones en estadística, incluidas las pruebas de hipótesis. Los parámetros a definir son:

- **X** es el valor cuya distribución desea obtener ($x_i = a$).
- **Media** es la media aritmética de la distribución (parámetro μ).
- **Desv_estándar** es la desviación estándar de la distribución (parámetro σ).
- **Acum** es un valor lógico que determina la forma de la función. Si el argumento Acum es VERDADERO, la función DISTR.NORM devuelve la función de distribución acumulada, es decir, la fórmula es el entero desde el infinito negativo a x de la fórmula dada.; si es FALSO, devuelve la función de masa de probabilidad.

- **Ejemplo:**

Calcular la probabilidad de que un estudiante de primero de ADE, tenga una altura superior a 175, sabiendo que la variable $X \sim \text{Altura de los estudiantes de ADE} \sim N(170, 20)$

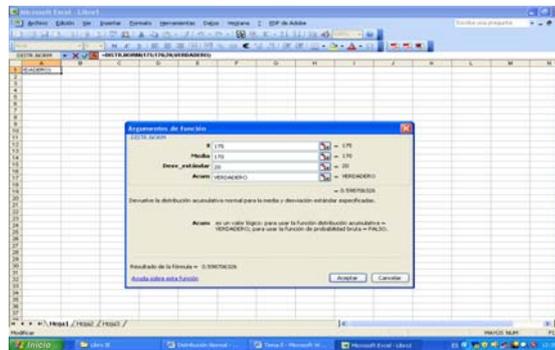


Imagen 7. Pantalla de Microsoft Excel con los argumentos y resultado
Fuente: Elaboración propia

La probabilidad de que un estudiante tenga una altura superior a 175, es del 59,87%.

4.2.1.2. Variable Normal Tipificada, $N(0, 1)$

La función **DISTR.NORMAL.ESTANDAR** nos permite calcular la probabilidad que, en una distribución normal de media cero y desviación típica uno, se encuentra por debajo del valor "a", ($P(Z \leq a) = ?$). Para ello, en insertar función, selecciona dentro del tipo de funciones estadísticas, el comando "**DISTR.NORM.ESTAND(z)**".

Esta opción devuelve la función de distribución normal estándar acumulativa. La distribución tiene una media de 0 (cero) y una desviación estándar de uno. Esta función nos proporciona las mismas probabilidades que la tabla estándar de áreas de curvas normales.

- **Ejemplo:**

Calcular la probabilidad de que una variable $N(0,1)$ tome un valor superior a 2.

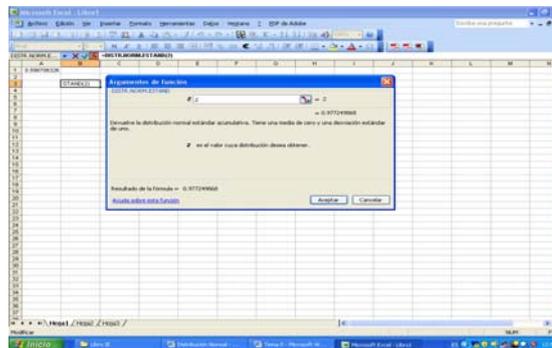


Imagen 8. Pantalla de Microsoft Excel con los argumentos y resultado
Fuente: Elaboración propia

4.2.1.3. Proceso de Normalización

Finalmente, si elegimos la opción “**NORMALIZACION**” obtendremos el valor normalizado de una distribución caracterizada por los argumentos media y desv_estándar.

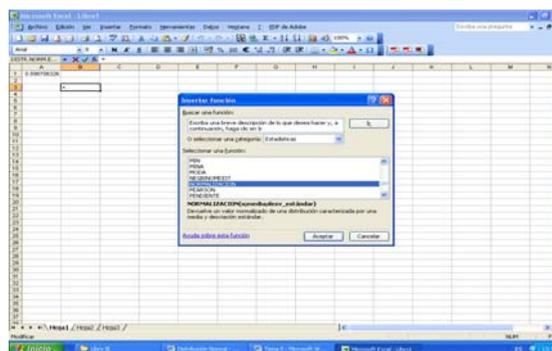


Imagen 9. Pantalla de Microsoft Excel para seleccionar la función
Fuente: Elaboración propia

Los parámetros a definir son:

- **X** es el valor que se desea normalizar.
- **Media** es la media aritmética de la distribución (parámetro μ).
- **Desv_estándar** es la desviación estándar de la distribución (parámetro σ).

• **Ejemplo:**

Volvamos a calcular el ejemplo del “Caso 1”, mediante esta opción. Calcular la probabilidad de que en estudiante de primero de ADE, tenga una altura superior a 175, sabiendo que la variable $X \sim \text{Altura de los estudiantes de ADE} \sim N(170, 20)$ ”

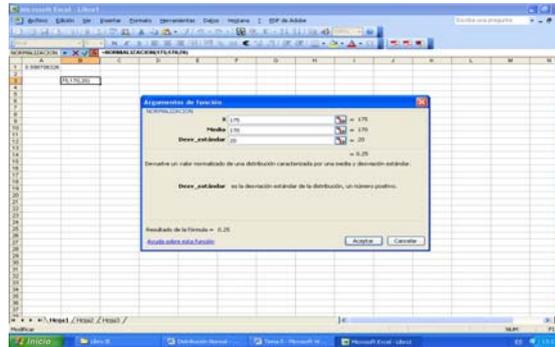


Imagen 10. Pantalla de Microsoft Excel para introducir los argumentos
Fuente: Elaboración propia

En este caso, el valor de 0,25 obtenido nos está indicando es que,

$$P(N(170,20) > 175) = P(N(0,1) > \frac{175 - 170}{20}) = P(N(0,1) > 0,25)$$

5. Cierre

En este objeto de aprendizaje hemos visto la utilización de Microsoft Excel para la estimación de probabilidades. Es cierto que actualmente existe una gran variedad de programas informáticos que pueden ser de gran utilidad para el cálculo de probabilidades, pero ¡no todos son tan accesibles y versátiles como Microsoft Excel!

Recuerda que hemos desarrollado, con ejemplos, el cálculo de la función de probabilidad y distribución para variables aleatorias discretas que siguen una distribución Binomial y/o de Poisson.

Así mismo, hemos desarrollado con ejemplos, el cálculo de la función de densidad y distribución para variables aleatorias continuas que siguen una distribución Normal.



6. Bibliografía

6.1. Libros:

- [1] Anderson, D.R.; Sweeney, D.J.; Williams, T.A. (2005). *Estadística para administración y economía*. (8ª Ed.). International Thomson Editores, México. ISBN 970-686-276-1
- [2] Arnaldos, F.; Díaz, M.T.; Faura, U.; Molera, L. y Parra, I. (2003). *Estadística descriptiva para economía y administración de empresas: cuestiones tipo test y ejercicios con Microsoft® Excel*. Madrid: AC.
- [3] Martín Pliego, F.J. (2004). *Introducción a la Estadística Económica y Empresarial*. (Ed.) Thomson. Madrid.
- [4] Mendenhall, W.; Reinmuth, J.E. (1978). *Estadística para administración y economía*. (Ed.) Grupo Editorial Iberoamericana. ISBN 968-7270-13-6.
- [5] Parra Frutos, I. (2003). *Estadística empresarial con Microsoft® Excel: Problemas de inferencia*. Segunda edición. Madrid: AC.
- [6] Piñole, R.; Moreno, A. y Caballero, A. (2002). *Análisis de datos y probabilidad. Excel como instrumento de cálculo*. Madrid: Civitas.
- [7] Uña Juárez, I; Tomero, V; San Martín, J. (2003). *Lecciones de cálculo de probabilidades*. (Ed.) Thomson Editores Spain. ISBN 84-9732-193-6.

6.2. Referencias de fuentes electrónicas:

- [8] http://www.uv.es/asepuma/XII/comunica/arnaldos_diaz_faura_molera_parra.pdf
- [9] http://www.uam.es/personal_pdi/economicas/mgizquie/excel_probabilidad.pdf