Utilización Práctica del

Teorema Central del Límite

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 conocer las características básicas del Teorema Central del Límite y sus posibles aplicaciones prácticas con la finalidad de construir una especie de catálogo al que acudir para el cálculo de probabilidades de distribuciones discretas, como binomial o Poisson mediante aproximaciones a la distribución normal.

2. Introducción

¿Qué diferencia existe en el cálculo de probabilidades de la variables aleatorias (V.A.) discretas, que siguen distribuciones binomiales o de Poisson y el cálculo de probabilidades de variables aleatorias continuas que siguen una distribución normal?

Una variable aleatoria se define continua cuando el conjunto de valores que puede tomar es un infinito continuo, es decir, puede tomar cualquier valor en un intervalo. Por el contrario, se define discreta cuando están medidas finitas o infinitas numerables, representan algo que podemos contar, y no suelen llevar decimales. Las distribuciones de probabilidad más utilizadas en variables discretas son la distribución binomial y la distribución de Poisson. La distribución más frecuente en el caso de las variables continuas es la distribución normal.

El Teorema Central del Límite indica que, en condiciones muy generales, la distribución de la suma de variables aleatorias tiende a una distribución normal cuando la cantidad de variables es muy grande. Es decir, garantiza una distribución normal cuando n es suficientemente grande.

En este objeto de aprendizaje, conoceremos las características y propiedades del Teorema Central del Límite. Utilizamos ejemplos y ejercicios donde descubriremos las razones por las cuales, en muchos campos de aplicación, se encuentran en todo momento distribuciones normales, o casi normales ya que cuando n es suficientemente grande, de variables independientes y todas ellas siguen el mismo modelo de distribución (cualquiera que éste sea), la suma de ellas se distribuye según una distribución normal, siendo aplica tanto a suma de variables discretas como de variables continuas, lo que facilitará el cálculo de las probabilidades. Finalmente, resaltamos los conceptos básicos de aprendizaje con respecto al Teorema Central del Límite y sus aplicaciones prácticas.

3. Objetivos

- Mostrar las características y propiedades del teorema Central del Límites

- Utilizar la distribución normal para aproximar el cálculo de probabilidades de variables discretas de Poisson.
- Utilizar la distribución normal para aproximar el cálculo de probabilidades de variables discretas de Binomial.

4. El Teorema Central de Límite

Este teorema afirma que la distribución de medias muéstrales tiende hacia una distribución normal, aunque las muestras procedan de una distribución no normal determinar un modelo de probabilidad para describir el comportamiento de una variable continua.

Es un Teorema de gran importancia en Estadística, especialmente para la parte de Inferencia Estadística. Establece que si $X_1,\ldots\ldots\ldots,X_n$ son variables aleatorias independientes con media μ_i y varianza σ_i^2, al margen del tipo de distribución que sigan los sumandos, la suma de todas ellas, $Y = X_1+\ldots\ldots+X_n$ tiende a distribuirse aproximadamente normal, con media $\mu = (\mu_1+\ldots\ldots+ \mu_n)$ y varianza $\sigma^2 =(\sigma_1^2+\ldots\ldots+\sigma_k^2)/n$, siendo las aproximaciones mejores a medida que aumenta n.

De acuerdo con el resultado siempre que observemos una variable que sea el resultado de muchas causas independientes, esperamos que su distribución sea aproximadamente normal, de ahí la frecuencia con la que se presentan en realidad variables aleatorias cuya distribución puede aproximarse a la distribución normal, debido a que muchas de éstas variables reales pueden considerarse la suma de variables independientes. Por ejemplo, las medidas físicas de una persona son debidas a muchas variables: genética, alimentación, ejercicio físico, etc. y se esperan que sigan una distribución normal.

Dado que la variable Binomial, no es más que la suma de n variables independientes (o la suma de los resultados obtenidos al efectuar n repeticiones de un experimento aleatorio con una variable aleatoria que sólo podía tomar dos posibles valores), su distribución tienda a aproximarse a la normal a medida que aumenta n, con media, la esperanza $E(X)=np$ y varianza $\sigma^2 = np(1-p)$, resultado demostrado por De Moivre en 1733. Este autor encontró que si X es una variable $B(n,p)$, la distribución:

$$\frac{X-np}{\sqrt{np(1-p)}}$$

Ecuación 1.Ecuación de convergencia de la distribución binomial a la distribución normal.

converge hacia una distribución normal tipificada o estandarizada, *N(0,1)* (Peña, 2001).

Existe mucha controversia para determinar el tamaño de muestra adecuado para que la aproximación a la normal proporcione resultados aceptables. Aunque muchos autores suele dar por aceptable cuando el producto *np(1-p)>5*, nosotros por mayor seguridad y fiabilidad de que las aproximaciones son adecuadas, trabajaremos con la regla de que ducho producto sea *np(1-p)≥* 9. En nuestro caso una variable Binomial *B(n,p)* se aproxima a una normal $N(\mu, \sigma)$, mediante la siguiente expresión:

$$B(n,p) \sim N(np, \sqrt{np(1-p)}$$

Ecuación 2. Ecuación de transformación distribución binomial a la distribución normal.

Esta misma situación aparece con variables de Poisson. Una variable aleatoria *X* que sigue una distribución *Ps(λ)*, puede aproximarse a una distribución normal con media, la esperanza *E(X)=* λ y varianza σ^2 = λ, cuando n es grande, lo que requiere elevados valores de λ. Aunque, como en el caso de la binomial, muchos autores suelen dar por aceptable cuando el producto λ *>5*, nosotros por mayor seguridad y fiabilidad de que las aproximaciones son adecuadas, trabajaremos con la regla de λ ≥ 9.

El procedimiento es similar al caso anterior. Si *X* es una variable *Ps(λ)*, la distribución:

$$\frac{X-\lambda}{\sqrt{\lambda}}$$

Ecuación 3.Ecuación de convergencia de la distribución de Poisson a la distribución normal.

converge hacia una distribución normal con tipificada o estandarizada, *N(0,1)*.
Es decir, una variable *Ps(λ),* se aproxima a una normal $N(\mu, \sigma)$, mediante la siguiente expresión:

$$Ps(\lambda) \sim N(\lambda, \sqrt{\lambda})$$

Ecuación 4. Ecuación de transformación distribución Poisson a la distribución normal.

Veamos algunos ejemplos:

- **Ejemplo 1**: *Calcular la probabilidad obtener 200 puntos al lanzar 70 dados.*

  Sea la variable X ~ número de puntos obtenidos en los 70 lanzamientos ~ B (n=70, p=1/6). Consideramos cada uno de los 70 lanzamientos como variables independientes, con lo cual podemos aplicar el Teorema Central del Límite y aproximarla a una distribución normal, pues se cumple que *np(1-p)= 9,72* es decir, ≥ 9.

X ~ número de puntos obtenidos en los 20 lanzamientos ~ B (n=20, p=1/6) ~ N(*E(X), σ(X)*), donde E(X) y *σ(X)* son la esperanza y la desviación típica de la distribución binomial, lanzamiento de 70 dados.

Aprovecharemos la propiedad, de que conocidos lo la esperanza (o puntos medios obtenidos al lanzar un dado) y la varianza de la variable puntos obtenidos al lanzar un dado, que sigue una distribución B (n=1, p=1/6), con una media $E_1(X)$ y una varianza $\sigma_1^2(X)$, podemos calcular, E(X) y $\sigma^2(X)$ como:

$$E(X)=70* E_1(X)=70*3{,}5=245$$

$$\sigma^2(X)=70* \sigma_1^2(X)=70*2{,}29= 160{,}3$$

Luego X~ número de puntos obtenidos en los 70 lanzamientos~ B (n=20, p=1/6) ~ N(*245, 12,66*)

$$P(X > 200) = P(N(0{,}1) > \frac{200-245}{12{,}66}) = 1 - P(N(0{,}1 > 3{,}55) = 1 - 0{,}00019 = 0{,}99981$$

- **Ejemplo 2**: *En un proceso de fabricación se producen en promedio 2 defectos por minuto. Calcular la aproximadamente la probabilidad de que en una hora se produzcan más de 150 defectos.*

X ~ Número defectos por minuto sigue una distribución de Poisson de parámetro λ=2.

En consecuencia, Y~ Número de defectos por hora sigue una distribución de Poisson de parámetro λ_{hora}=60*2=120.

Consideramos cada uno de los defectos por minuto como variables independientes, con lo cual podemos aplicar el Teorema Central del Límite y aproximarla a una distribución normal, pues se cumple que $\lambda_{hora} \geq 9$.

Y~ Número de defectos por hora sigue una distribución de Ps(λ_{hora}120) ~N(*E(X), σ(X)*), donde E(X) y *σ(X)* son la esperanza y la desviación típica de la de Poisson.

$$E(X)= \lambda_{hora}=120$$

$$\sigma^2(X)= \lambda_{hora}=120$$

$$P(Y > 150) = P(N(0{,}1) > \frac{200-120}{10{,}95}) = P(N(0{,}1 > 2{,}74) = 0{,}0031$$

5. Cierre

El Teorema Central del Límite establece que la suma de n variables aleatorias independientes de varianza finita e idéntica distribución tiende a la **distribución normal** cuando n tiende a infinito.

El Teorema Central del Límite, **permite calcular razonablemente bien las probabilidades de variables que siguen una distribución Binomial y de Poisson**, siempre que el tamaño de muestra sea suficientemente grande, lo cual equivale a que se cumpla que $np(1-p)\geq 9$ o que $\lambda\geq 9$, respectivamente.

Una variable Binomial $B(n,p)$ se aproxima a una normal $N(\mu, \sigma)$, mediante la siguiente expresión:

$$B(n,p) \sim N(np, \sqrt{np(1-p)})$$

Una variable Poisson $Ps(\lambda)$ se aproxima a una normal $N(\mu, \sigma)$, mediante la siguiente expresión:

$$Ps(\lambda) \sim N(\lambda, \sqrt{\lambda})$$

6. Bibliografía

6.1. Libros:

[1] Martín Pliego, F.J. (2004). *Introducción a la Estadística Económica y Empresarial*. (Ed.) Thomson. Madrid.
[2] Mendenhall, W.; Reinmuth, J.E. (1978). *Estadística para administración y economía*. (Ed.) Grupo Editorial Iberoamericana. ISBN 968-7270-13-6.
[3] Peña, D. (2001). *Fundamentos de Estadística*. (Ed.) Alianza Editorial, S.A. Madrid. ISBN: 84-206-8696-4.
[6] Romero, R y Zúnica, L.R. (1993). Estadística (Proyecto de Innovación Educativa). SPUPV-93.637.
[7] Romero, R y Zúnica, L.R. (2000). Introducción a la Estadística. (Ed.). SPUPV- 2000.4071.

6.2. Referencias de fuentes electrónicas:

[8] http://personal5.iddeo.es/ztt/Tem/t21_distribucion_normal.htm (Consultado 17/10/08).
[9] http://74.125.39.104/search?q=cache:nIXKW90gaC0J:www.udl.es/usuaris/seio2003/treballs/06_21.pdf+PAPEL+PROBABIL%C3%8DSTICO&hl=es&ct=clnk&cd=7&gl=es
[10] http://sauce.pntic.mec.es/~jpeo0002/Archivos/PDF/XT03.pdf
[11] http://www.vitutor.com/pro/5/a_g.html
[12] http://es.geocities.com/pilar_zutabe/EJERCICIOS/1BACHILLERHUMANISTICO/Ejerciciosdistribucionnormal.htm
[13] http://www.digeo.cl/asignaturas/mat/Ejercicios-Distribucion-Normal.pdf
[14] http://www.fisterra.com/mbe/investiga/distr_normal/distr_normal.asp
[15] www1.uprh.edu/.../La%20distribucion%20normal/Modulo%20Sobre%20La%20Distribucion%20Normal%