Subject: Probability and Statistics

Class VII: Confidence Intervals and Hypothesis Testing

UdeA

Bioengineering

Francisco José Campuzano Cardona

Bioengineerer, MSc in Engineering

UdeA

Suponga que X es un variable aleatoria con una distribución cualquiera.

Sea: $\boldsymbol{\mu}_x$ = la media de **X**, y $\boldsymbol{\sigma}_x$ = la desviación estándar de **X**.

Si se toman muestras aleatorias de X de tamaño n, múltiples veces, con n grande, y a cada muestra se le calcula la media, entonces tenemos una nueva variable aleatoria $\bar{X}$ que tenderá a tener una distribución normal:

$$\bar{X} \sim N\left(\boldsymbol{\mu}_x, \frac{\boldsymbol{\sigma}_x}{\sqrt{n}}\right)$$

UdeA

Suponga que X es un variable aleatoria con una distribución cualquiera.

Sea: $\boldsymbol{\mu}_x$ = la media de **X**, y $\boldsymbol{\sigma}_x$ = la desviación estándar de **X**.

Si se toman muestras aleatorias de X de tamaño n, múltiples veces, con n grande, y se hace la sumatoria de cada muestra, entonces tenemos una nueva variable aleatoria ΣX que tenderá a tener una distribución normal:

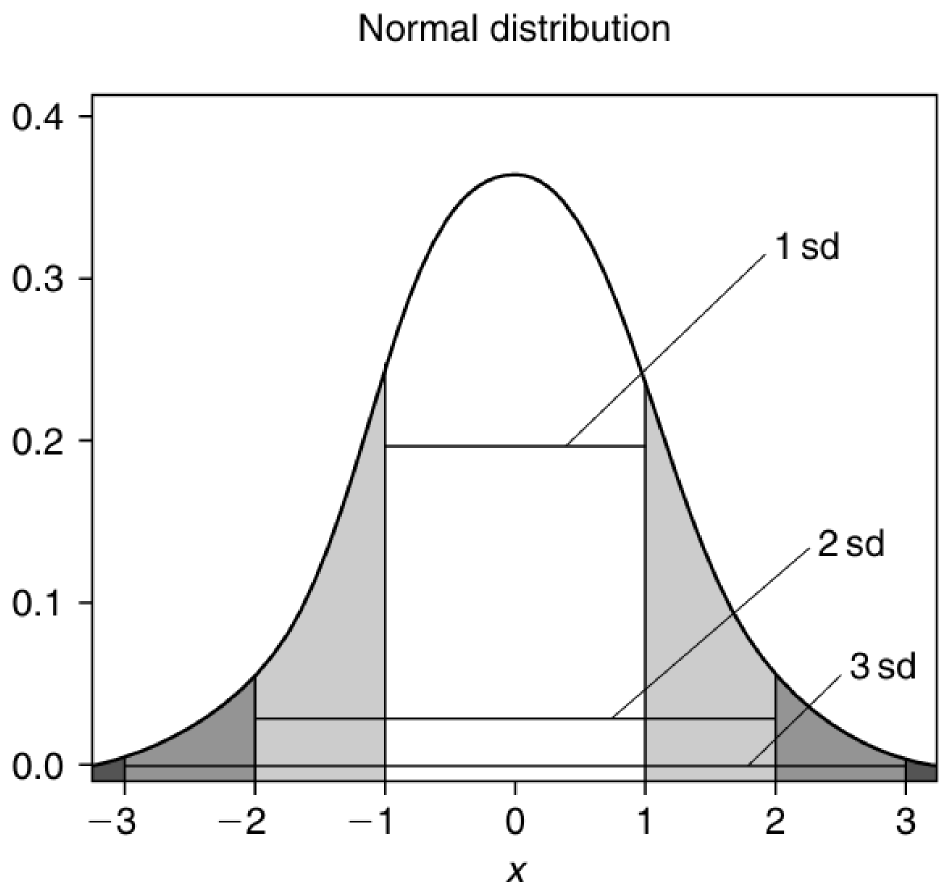
$$\Sigma X \sim N\left(n\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x\sqrt{n}\right)$$

Confidence Intervals

Como habíamos dicho antes, el muestro permite realizar estimación de parámetros de una población. Pero esta estimación puede tener un error, entonces el estadístico (estimador del parámetro) puede distribuirse en un rango de valores conocido como **Intervalo de Confianza.**

Entonces cuando realizamos un muestro, y determinamos la media y la desviación estándar de la muestra, recordamos que si repetimos este proceso muchas veces tendríamos una distribución normal:

$$\bar{X} \sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right)$$



UdeA

Las pruebas de hipótesis, tienen como propósito ayudar a determinar si el azar podría ser responsable de un efecto observado.

En algunos casos se quiere comparar una muestra con un valor de referencia que puede ser un parámetro supuesto, o una distribución. **Se llaman Pruebas de una muestra**

En otros casos se tienen dos grupos que se quieren que sean comparados para determinar si existen diferencias, para eso se realizan **pruebas de comparación de dos grupos.** Los grupos a comparar difieren en algo.

UdeA

Ejemplos:

- Un hospital quiere determinar si el nivel medio de colesterol en un grupo de pacientes con riesgo cardiovascular es diferente al valor normal de referencia en la población general (200 mg/dL).

- Se realiza un estudio clínico para determinar si un medicamento es mejor que otro medicamento estándar en el tratamiento de una enfermedad. En este caso se debe determinar si el efecto terapéutico del medicamento nuevo, es mejor que el del control, para esto un grupo debe ser sometido al tratamiento convencional, y otro grupo al tratamiento con el nuevo medicamento.

- Se quiere determinar si adicionar nanofibras de celulosa a una membrana de PVA mejora la resistencia a la tracción de la membrana. Para esto debe tenerse un grupo de membranas suplementadas con nanofibras de celulosa (ese es el tratamiento) y otro grupo de membranas que no fueron adicionadas con nanofibras, las cuales serían el control (no fueron sometidas al tratamiento).

En las pruebas de hipótesis siempre se plantean dos hipótesis que deben dar cuenta de todas las posibilidades del experimento.

"Dada la tendencia humana a reaccionar ante comportamientos inusuales pero aleatorios e interpretarlos como algo significativo y real, en nuestros experimentos requeriremos prueba de que la diferencia entre los grupos es más extrema de lo que el azar podría producir razonablemente"

Hipótesis Nula (H0): Es la hipótesis que sugiere que el efecto observado es responsabilidad del azar.

Hipótesis Alternativa (H1): generalmente es lo que se desea probar, y es lo contrario a la hipótesis nula.

UdeA

Conceptos importantes en las pruebas de hipótesis

Valor-p: es la probabilidad de obtener valores tan inusuales o extremos como los valores observados. Es decir la probabilidad de que lo que se observó sea cuestión del azar. Este valor se determina según la prueba usada. Esta valor está entre 0 y 1.

Alfa: Es el umbral de probabilidad máximo que se acepta para la probabilidad de que lo que se observó sea cuestión del azar. Valores típicos 5% o 1% (0.05 o 0.01), es el complemento del intervalo de confianza.

Error de Tipo I: Es el error asociado a afirmar que un efecto es real, cuando en realidad es resultado del azar.

Error de Tipo II: Es el error asociado a afirmar que un efecto es resultado del azar, cuando en realidad es un efecto real.

UdeA

Interpretación del valor-p

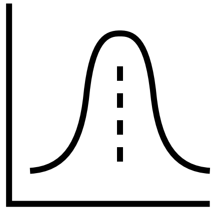
El **valor-p** se encuentra entre 0 y 1. El valor **alfa** es un umbral elegido arbitrariamente, usualmente 0.05 (5%) o 0.01 (1%). El **valor-p** siempre es comparado con el valor **alfa** elegido, para determinar la validez de las hipótesis.

Cuando el **valor-p es menor al valor alfa**, entonces se **rechaza H0**, y se comprueba H1. En caso contrario, si el valor-p es mayor a alfa, se debe rechazar H1 y aceptar H0

El error de tipo I se asocia al primer caso, y el de tipo II al segundo. El error de tipo II, puede darse por una muestra insuficiente, entonces en realidad el rechazo de H1 es más bien un caso de “efecto no probado”, aumentar el n podría cambiar el resultado.

Las pruebas estadísticas también pueden clasificarse en otros dos grupos, pruebas **paramétricas** y pruebas **no paramétricas.**

En el primer caso, las pruebas paramétricas, parten de suponer cierto comportamiento en los datos, que debe ser comprobado para que las pruebas tengan validez.



En el caso de las pruebas no paramétricas, estos supuestos son menos estrictos o inexistentes.

Tests types

UdeA

Criterio	Pruebas Paramétricas	Pruebas No Paramétricas
Suposiciones	Requieren normalidad y homogeneidad de varianza	No requieren normalidad
Ejemplos de pruebas	T-test, ANOVA	Mann-Whitney, Kruskal-Wallis
Datos	Cuantitativos (continuos)	Cuantitativos (ordinales) o con valores atípicos
Eficiencia	Más potentes si se cumplen las suposiciones	Menos potentes pero más robustas
Tamaño de muestra	Necesitan muestras más grandes	Funcionan bien con muestras pequeñas

UdeA

En muchas ocasiones se desea comparar una muestra, con un valor de referencia, o una población estándar de referencia.

Prueba t de una sola muestra:
Sirve para comparar si la media de una muestra es diferente o no de un valor medio de referencia. **Tiene el supuesto de que los datos son normales.**

H0: *La media de la muestra es igual al valor poblacional esperado.*
H1: *La media de la muestra es diferente al valor esperado.*

Ejemplo: ¿El nivel medio de glucosa en una muestra de pacientes es diferente a 100 mg/dL?

UdeA

Prueba de una proporción (Z para proporciones)

Sirve para comparar la proporción de una característica en una muestra con un valor poblacional conocido. **Tiene el supuesto de que los datos son binomiales**. Por el teorema del central del limite, si n grande se aproxima a la normal si $np \geq 5$ y $n(1-p) \geq 5$. Con p la proporción que se está verificando.

H_0: *La proporción de la muestra es igual a la proporción de la población.*
H_1: *La proporción de la muestra es diferente a la proporción de la población.*

Ejemplo: Un investigador quiere evaluar si la tasa de pacientes con hipertensión en su hospital es mayor al 30% reportado en estudios nacionales.

UdeA

Prueba de una proporción (Z para proporciones)

La prueba se puede hacer bilateralmente o unilateralmente.

En el caso bilateral, importa saber si la proporción es diferente sin importar si es mayor o menor. En el caso unilateral, importa saber si la proporción es mayor, o menor.

UdeA

Prueba de una proporción (Z para proporciones)
Ejemplos.

Queremos saber si un nuevo material biodegradable tiene una tasa de biodegradación menor que el estándar del 90%.
$H0{:}p=0.9$, $H1{:}p<0.9$

Un sensor vestible detecta deshidratación con una precisión esperada de 80%. Queremos saber si un nuevo algoritmo tiene una precisión diferente, ya sea mayor o menor.

$H0{:}p=0.8$, $H1{:}p\neq0.8$

UdeA

Pruebas de Normalidad

La normalidad se puede analizar a través de pruebas de hipótesis, en este caso se compara una muestra, con una distribución estándar, la normal.

- Kolmogorov-Smirnoff (n>50)
- Shapiro-Wilks (n<50)

Ho: Los datos provienen de una distribución normal
H1: Los datos no provienen de una distribución normal

Como la normalidad es un supuesto de muchas pruebas paramétricas, se usa estas pruebas para probarlo.

UdeA

Ejercicio

En la carpeta correspondiente a la clase de hoy, encontrará un enunciado para el ejercicio entregable del día de hoy.

Subject: Probability and Statistics

UdeA

Bioengineering

¡Thanks!

Francisco José Campuzano Cardona

Bioengineering. MSc in Engineering