# Estadística Inferencial

Clifford Torres

Facultad de Ciencias (UNI)

September 16, 2019

# Ejemplos: Selección del tamaño muestral

- La reacción de un individuo a un estímulo en un experimento psicológico puede tomar una de dos formas, A o B. Si un experimentador desea estimar la probabilidad p de que una persona reaccione en una forma A, cuántas personas deben incluirse en el experimento? Suponga que el experimentador estará satisfecho si el error de estimación es menor que .04 con probabilidad igual a .90. Suponga también que él espera que p se encuentre en algún punto cercano a 0.6.

# Ejemplos: Selección del tamaño muestral

- Un experimentador desea comparar la efectividad de dos métodos de capacitación para obreros que van a realizar una operación de ensamble. Los obreros seleccionados han de dividirse en dos grupos de igual tamaño, el primero para recibir el método 1 de capacitación y el segundo el método 2 de capacitación. Después de la capacitación cada obrero realizará la operación de ensamble y se registrará el tiempo que le tome hacerlo. El experimentador espera que las mediciones para ambos grupos tengan una amplitud de aproximadamente 8 minutos. Si la estimación de la diferencia en los tiempos promedio de ensamble debe ser correcta con una variación de no más de 1 minuto con probabilidad .95, cuántos trabajadores deben incluirse en cada grupo de capacitación?

# Distribución de probabilidad Gamma

## Definición

*Se dice que una variable aleatoria $Y$ tiene una distribución gamma con párametros $\alpha > 0$ y $\beta > 0$ si y solo si la función de densidad de $Y$ es*

$$f(y) = \begin{cases} \frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)} & , \quad 0 \leq y < \infty \\ 0 & , \quad c.c. \end{cases}$$

*donde*

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{y}dy.$$

## Ejercicio

*Probar que*

$$E[Y] = \alpha\beta \quad y \quad V[Y] = \alpha\beta^2.$$

# Distribución de probabilidad Gamma

**Definición**

*Sea* $v$ *un entero positivo. Se dice que una variable aleatoria* $Y$ *tiene una distribución ji cuadrada* $(\chi^2)$ *con* $v$ *grados de libertad si y solo si* $Y$ *es una variable aleatoria con distribución gamma y parámetros* $\alpha = v/2$ *y* $\beta = 2$.

**Ejercicio**

*Probar que*

$$E[Y] = v \quad y \quad V[Y] = 2v.$$

# Distribución de probabilidad Gamma

**Definición**

*Se dice que una variable aleatoria $Y$ tiene una distribución exponencial con parámetro $\beta$ si y solo si $Y$ es una variable aleatoria con distribución gamma y parámetros $\alpha = 1$ y $\beta > 0$.*

**Ejercicio**

*Probar que*

$$E[Y] = \beta \quad y \quad V[Y] = \beta^2.$$

# Distribuciones muestrales relacionadas con $N(\mu, \sigma^2)$

**Teorema**

*Sea $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria de tamaño $n$ de una distribución normal con media $\mu$ y varianza $\sigma^2$. Entonces $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ está distribuida normalmente con media $\mu_{\bar{Y}} = \mu$ y su varianza $\sigma^2_{\bar{Y}} = \sigma^2/n$.*

Se deduce que

$$Z = \frac{\bar{Y} - \mu_{\bar{Y}}}{\sigma_{\bar{Y}}} = \sqrt{n}\frac{\bar{Y} - \mu}{\sigma} \sim N(0,1).$$

# Distribuciones muestrales relacionadas con $N(\mu, \sigma^2)$

**Teorema**

*Si $Y_1, Y_2, \cdots, Y_n$ está definida como en el Teorema previo. Entonces $Z_i = (Y_i - \mu)/\sigma$ son variables aleatorias normales estándar e independientes, $i = 1, 2, \cdots, n$, y $\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} (\frac{Y_i - \mu}{\sigma})^2$ tienen una distribución $\chi^2$ con $n$ grados de libertad (gl).*

**Teorema**

*Sea $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria de una distribución normal con media $\mu$ y varianza $\sigma^2$. Entonces*

$$\frac{(n-1)s^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

*tiene una distribución $\chi^2$ con $(n-1)$ gl. También, $\bar{Y}$ y $s^2$ son variables aleatorias independientes.*

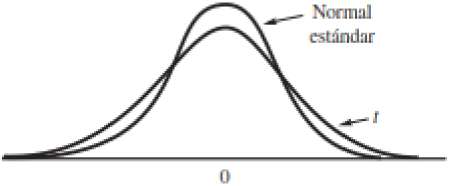
# Distribuciones muestrales relacionadas con $N(\mu, \sigma^2)$

**Definición**

*Sea $Z$ una variable aleatoria normal estándar y sea $W$ una variable con distribución $\chi^2$ con $v$ grados de libertad. Entonces, si $W$ y $Z$ son independientes,*

$$T = \frac{Z}{\sqrt{W/v}}$$

*se dice que tiene una distribución $t$ (t-student) con $v$ grados de libertad.*




**Ejercicio**

*Probar que:* $v > 1: \; E[T] = 0 \quad y \quad v > 2: \; V[T] = v/(v-2)$

# Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- Los intervalos de confianza para una media poblacional $\mu$, que estudiaremos, están basados en la suposición de que la muestra del experimentador se ha seleccionado aleatoriamente de entre una población normal.
- Los intervalos son apropiados para muestras de cualquier tamaño y los coeficientes de confianza de los intervalos son cercanos a los valores especificados aun cuando la población no sea normal, mientras la desviación no sea excesiva.
- Raras veces conocemos la forma de la distribución de frecuencia poblacional antes de muestrear y, en consecuencia, si un estimador de intervalo debe ser de cualquier valor, debe funcionar razonablemente bien aun cuando la población no sea normal.

# Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- **Funcionar bien** significa que el coeficiente de confianza no debe ser afectado por desviaciones pequeñas de la normalidad.
- Para la mayor parte de las distribuciones poblacionales en forma de campana, los estudios experimentales indican que estos intervalos de confianza mantienen coeficientes de confianza cercanos a los valores nominales empleados en su cálculo.

## Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- Supongamos que $Y_1, Y_2, \cdots, Y_n$ representa una muestra aleatoria seleccionada de una población normal tal que $\bar{Y}$ y $S^2$ representa la media y la varianza muestral.
- Nos gustaría construir un intervalo de confianza para la media poblacional, cuando $V(Y_i) = \sigma^2$ sea desconocida y el tamaño de la muestra sea demasiado pequeño para permitirnos aplicar las técnicas de muestra grande expuestas anteriormente.

## Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- Dadas las suposiciones que acabamos de indicar

$$T = \frac{Y - \mu}{s/\sqrt{n}}$$

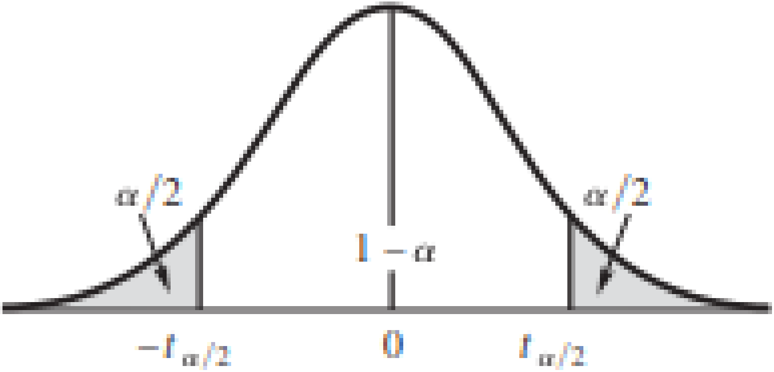
tiene una distribución $t$ con $(n-1)$ grados de libertad.

- La cantidad $T$ sirve como cantidad pivote que usaremos para formar un intervalo de confianza para $\mu$. Asimsimo, pode-mos hallar valores $t_{\alpha/2}$ y $-t_{\alpha/2}$ de modo que

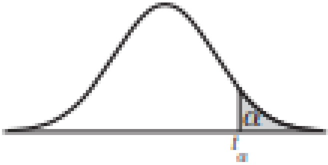
$$P(-t_{\alpha/2} \leq T \leq t_{\alpha/2}) = 1 - \alpha$$

- La distribución $t$ tiene una función de densidad muy semejante a la densidad normal estándar excepto que las colas son más anchas.
- Recuerde que los valores de $t_{\alpha/2}$ dependen de los grados de libertad $(n-1)$ así como del coeficiente de confianza (1 - $\alpha$).

# Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$



# Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$



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

## Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- El intervalo de confianza para $\mu$ se obtiene al manipular las desigualdades del enunciado de probabilidad, de modo análogo al empleado anteriormente. En este caso, el intervalo de confianza resultante para $\mu$ es

$$\bar{Y} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

- De acuerdo con las suposiciones anteriores, también podemos obtener límites de confianza unilaterales de $100(1-a)\%$ para $\mu$. Observe que $t_\alpha$, dada en la Tabla, es tal que

$$P(T \leq t_\alpha) = 1 - \alpha$$

- Sustituyendo $T$ en esta expresión y manipulando la desigualdad resultante obtenemos

$$P(\bar{Y} - t_\alpha(s/\sqrt{n}) \leq \mu) = 1 - \alpha$$

## Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- Entonces, $Y - t_\alpha(S/\sqrt{n})$ es un límite de confianza inferior de $100(1-\alpha)\%$para $\mu$.
- De la misma manera Y$-t_\alpha(S/\sqrt{n})$ es un límite de confianza superior de $100(1-\alpha)\%$ para $\mu$.
- Al igual que en el caso de una muestra grande, si determinamos los límites de confianza superior e inferior $100(1-a)\%$ para $\mu$ y usamos los límites respectivos como puntos extremos para un intervalo de confianzas el intervalo bilateral resultante tiene coeficiente de confianza igual a $1-2\alpha$.

## Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- Entonces, $Y - t_\alpha(S/\sqrt{n})$ es un límite de confianza inferior de $100(1-\alpha)\%$para $\mu$.
- De la misma manera $Y - t_\alpha(S/\sqrt{n})$ es un límite de confianza superior de $100(1-\alpha)\%$ para $\mu$.
- Al igual que en el caso de una muestra grande, si determinamos los límites de confianza superior e inferior $100(1-a)\%$ para $\mu$ y usamos los límites respectivos como puntos extremos para un intervalo de confianzas el intervalo bilateral resultante tiene coeficiente de confianza igual a $1-2\alpha$.

## Intervalos de confianza de una muestra pequeña para $\mu$ y $\mu_1 - \mu_2$

- Un fabricante ha inventado una nueva pólvora que fue probada en ocho proyectiles. Las velocidades resultantes en la boca del cañón, en pies por segundo, fueron las siguientes: 3005, 2925, 2935, 2965, 2995, 3005, 2937, 2905 Encuentre un intervalo de confianza de 95% para el verdadero promedio de velocidad $\mu$ para proyectiles de este tipo. Suponga que las velocidades en la boca del cañón están distribuidas normalmente en forma aproximada.