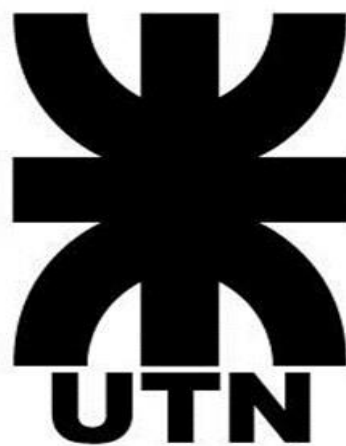


**Universidad Tecnológica Nacional**

Facultad Regional Córdoba

Año 2025



**Trabajo Práctico De Laboratorio**

**Simulación 4K1**

**Docentes:**

Prof. Berrotaran, Juan José

Prof. Ravera, Emanuel Mariano

**Integrantes:**

- Carrió, Tomás 94763
- Ferreyra Lammertyn, Franca 85341
- Fernández, Pablo Gabriel 95695
- Gon, Federico 94470
- Guardatti, Nicolás 82180
- Jorge Lorenzo, Francisco 86582
- Muñoz, Milena 82488
- Rinaudo, Martin 89615

## **Índice**

<b>Introducción.....</b>	<b>3</b>
<b>Consignas .....</b>	<b>4</b>
<b>Desarrollo.....</b>	<b>5</b>
<b>Conclusión .....</b>	<b>7</b>

## Introducción

En el presente trabajo práctico se llevará a cabo la toma y análisis de datos de un sistema real.

Las variables a presentar fueron tomadas de una fuente que estará documentada al final de la presentación.

En la resolución se especificará claramente la variable aleatoria medida y en qué unidad se ha expresado. Posteriormente, se determinará la distribución de probabilidad que mejor se ajuste a los datos, considerando el histograma de frecuencias.

Una vez identificada la distribución, se planteará hipotéticamente el comportamiento probabilístico de la muestra a través de la formulación de una Hipótesis Nula. Esta hipótesis será sometida a prueba mediante Chi Cuadrado y Kolmogorov-Smirnov, el cual será realizado con agrupamiento de intervalos si fuese necesario para un análisis más robusto.

Finalmente, se presentarán las conclusiones obtenidas a partir del análisis realizado, destacando los resultados más relevantes y las implicaciones que estos puedan tener en el contexto del sistema estudiado.

## Consignas

- a) Realizar la toma de datos de un sistema real (mínimo 200 datos).
- b) Determinar a que distribución de probabilidad se puede asimilar la misma.
- c) Realizar su histograma y gráfica.
- d) Determinar hipotéticamente el comportamiento probabilístico de la muestra.  
(Hipótesis Nula)
- e) Realizar las 2 pruebas que aceptan o rechazan tal hipótesis.
- f) Repetir esta experiencia con otra muestra diferente de datos (tamaño mínimo de la muestra 200).

## Desarrollo

La toma de datos de ambas variables se realizó desde la siguiente fuente:  
<https://www.kaggle.com/datasets/rodrigogastonrubio/futbolargentino?resource=download>

Como primera variable se ha tomado la edad de los jugadores de fútbol de los equipos de Arsenal, Independiente, Banfield, Argentinos Jr y Boca Juniors pertenecientes a la liga Argentina de futbol del año 2008.

Como segunda variable se tomó la altura (en metros) de los jugadores de futbol de Arsenal, Argentinos Jr, Boca y Colón pertenecientes a la liga Argentina de futbol del año 2008.

Una vez obtenidas las variables calculamos lo siguiente:

- N (tamaño de la muestra)
- Máximo
- Mínimo
- Rango
- Intervalos
- Amplitud
- Media
- Varianza muestra
- Varianza población
- Desviación muestra
- Desviación población

Después de haber hecho los cálculos procedemos a realizar la tabla para verificar a que tipo de distribución se asemejan los datos buscados.

Para la tabla tuvimos en cuenta lo siguiente:

- Intervalos
- Límite inferior
- Límite superior
- Frecuencia observada
- Frecuencia esperada

- Chi cuadrado

Ya calculada la frecuencia observada realizamos el histograma para poder determinar a que distribución pertenecen las variables. Como conclusión llegamos a que ambas se asemejan a una distribución normal.

Luego de determinar a que tipo de distribución pertenecen las variables calculamos la frecuencia esperada, en la que su fórmula dependerá a que tipo de distribución pertenecen las variables.

Como siguiente paso llevamos a cabo la prueba de bondad de Chi Cuadrado para comprobar si se acepta o se rechaza la hipótesis.

Para realizar la comprobación de la hipótesis en **Chi Cuadrado** primero realizamos la sumatoria de  $((FO - FE)^2/FE)$  (valor de Chi calculado) y luego buscamos el CHI de tabla usando la formula **+PRUEBA.CHI.INV( $\alpha$ ; intervalos - 1 - 2)**. Se le resta 2 ya que son 2 datos empíricos.

Como resultado, el valor de CHI calculado en ambos intervalos nos da que es mayor al CHI de tabla por ende se **rechazan ambas hipótesis nulas**.

Para realizar la comprobación de la hipótesis en **Kolmogorov-Smirnov** tenemos que calcular las probabilidades de la frecuencia observada con su acumulada (P.F.O y A.P.F.O) y las probabilidades de la frecuencia esperada (P.F.E) con su acumulada A.P.F.E). Cuando tenemos esto realizamos el cálculo  $(A.P.F.O - A.P.F.E)$  en valor absoluto.

El valor de K-S calculado va a ser el número máximo de la columna  $(A.P.F.O - A.P.F.E)$ .

Para calcular el valor de K-S aplicamos la formula  **$K\alpha/\sqrt{N}$** .

Como resultado, el valor de K-S calculado en la primera variable nos da mayor al K-S de tabla, por ende, **se rechaza hipótesis nula**, pero en nuestra segunda variable el valor de K-S calculado nos da menor que el K-S de tabla, por ende, se **acepta hipótesis nula**.

## Conclusión

En este trabajo, se llevó a cabo un análisis estadístico de dos variables relacionadas con el ámbito del deporte, específicamente la edad y la altura de jugadores de fútbol en la liga Argentina de futbol del año 2008.

A través de un proceso meticuloso de recolección y análisis de datos, se determinó que ambas variables siguen una distribución normal, lo cual se corroboró mediante el uso de histogramas de frecuencias y el cálculo de la frecuencia esperada.

Posteriormente, se formuló una hipótesis nula para cada variable, la cual fue sometida a prueba utilizando el test de Chi Cuadrado y de Kolmogorov-Smirnov. Los resultados obtenidos indicaron que, en ambos casos, el valor de Chi Cuadrado calculado fue mayor que el valor crítico de la tabla, lo que llevó al rechazo de las hipótesis nulas planteadas y que en Kolmogorov-Smirnov en la primera variable se rechaza hipótesis nula y en la segunda variable se acepta.