



Instituto Centroamericano
de Administración Pública

MAESTRÍA EN GERENCIA DE LA SALUD
CURSO: MÉTODOS ESTADÍSTICOS
PROYECTO FINAL

PROFESORA: MSC. SOFÍA ARBUROLA BRICEÑO

JOSÉ MANUEL PICADO CHANG

NOVIEMBRE 2021



Bases de datos utilizadas

1. Incidencia de tumores malignos en la población masculina y femenina, según localización anatómica y grupos de edad, Costa Rica, 2014.

Tomada de:

<https://www.ministeriodesalud.go.cr/index.php/vigilancia-de-la-salud/estadisticas-y-bases-de-datos/estadisticas/estadistica-de-cancer-registro-nacional-tumores/incidencia>

2. Cáncer de próstata. Tomada de:

<https://www.kaggle.com/sajidsaifi/prostate-cancer>



Estadísticas descriptivas de la incidencia de tumores malignos en la población masculina y femenina, según localización anatómica y grupos de edad, Costa Rica, 2014.

Masculino.

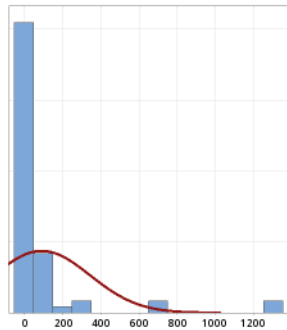
Variable	Conteo total	Media	Desv.Est.	Varianza	Mínimo	Q1	Mediana	Q3	Máximo	Rango	Modo	N para moda	Curtosis
TOTAL	62	78.2	221.1	48870.9	0.0	2.8	10.0	54.3	1399.0	1399.0	1	7	25.58

Femenino.

Variable	Conteo total	Media	Desv.Est.	Varianza	Mínimo	Q1	Mediana	Q3	Máximo	Rango	Modo	N para moda	Curtosis
TOTAL	66	91.8	252.2	63627.3	0.0	2.0	11.0	53.8	1320.0	1320.0	2	9	16.68



Informe de resumen de TOTAL



Prueba de normalidad de Anderson-Darling

A-cuadrado	15.58
Valor p	<0.005
Media	91.76
Desv.Est.	252.24
Varianza	63627.29
Asimetría	4.0298
Curtosis	16.6786
N	66

Mínimo	0.00
1er cuartil	2.00
Mediana	11.00
3er cuartil	53.75
Máximo	1320.00

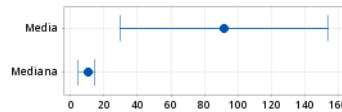
Intervalo de confianza de 95% para la media
29.75 153.77

Intervalo de confianza de 95% para la mediana
5.00 15.00

Intervalo de confianza de 95% para la desviación estándar
215.35 304.51

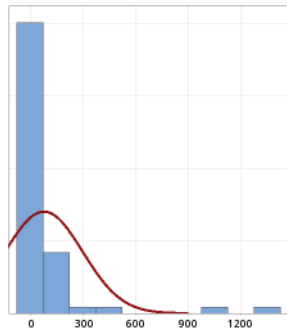


Intervalos de confianza de 95%



Masculino

Informe de resumen de TOTAL



Prueba de normalidad de Anderson-Darling

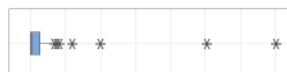
A-cuadrado	13.86
Valor p	<0.005
Media	78.23
Desv.Est.	221.07
Varianza	48870.93
Asimetría	4.8982
Curtosis	25.5776
N	62

Mínimo	0.00
1er cuartil	2.75
Mediana	10.00
3er cuartil	54.25
Máximo	1399.00

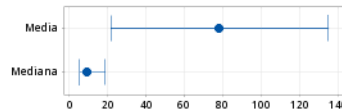
Intervalo de confianza de 95% para la media
22.09 134.37

Intervalo de confianza de 95% para la mediana
5.83 19.17

Intervalo de confianza de 95% para la desviación estándar
187.85 268.66



Intervalos de confianza de 95%



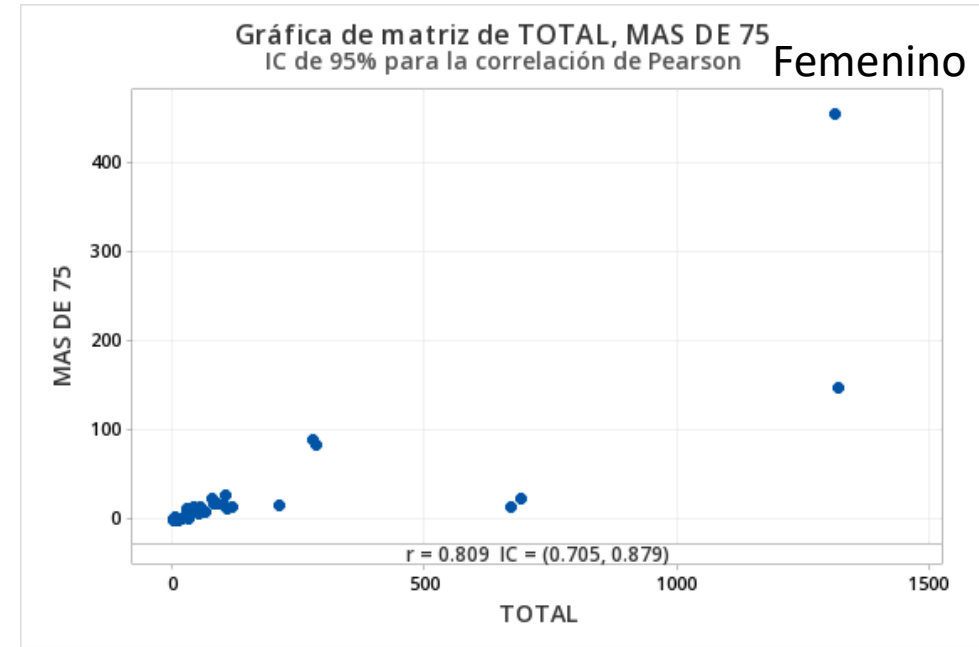
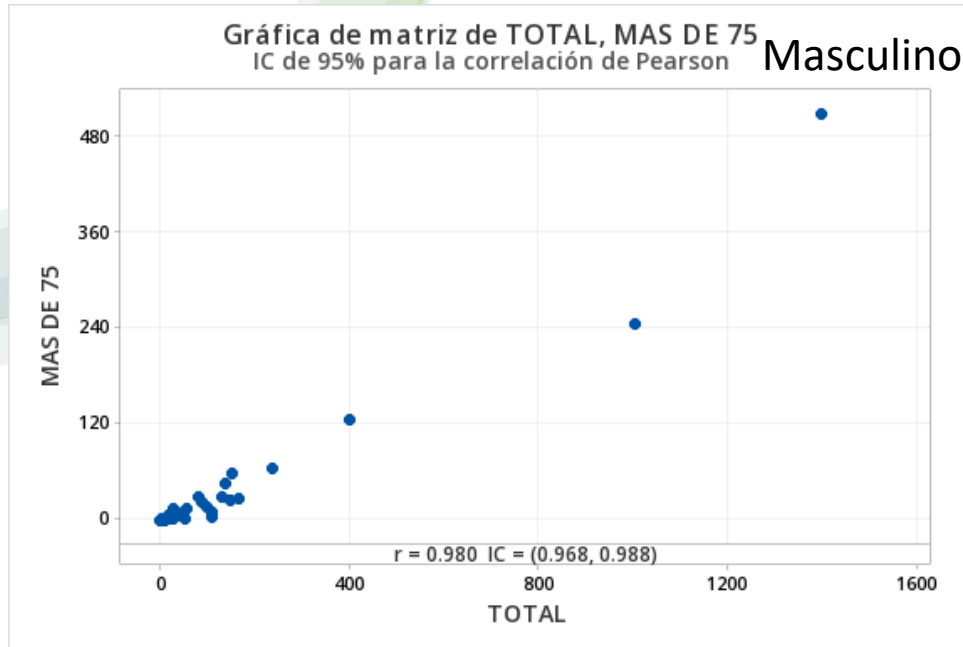
Femenino

Conclusiones.

1. En la MEDIA se documenta para el sexo masculino 78.2 casos y en el sexo femenino 91.8 casos, siendo estos valores que no representan de forma adecuada el centro de los datos.
2. Para la MEDIANA se documenta que el centro de los datos para el sexo masculino es de 10 casos y para el sexo femenino de 11 casos, valores muy similares.
3. La MODA se representa en 7 ocasiones la incidencia de 1 caso en el sexo masculino y en 9 ocasiones 2 casos en el sexo femenino.
4. El RANGO se observa para el sexo masculino y femenino que oscila desde los 0 casos en algunos tipos de cánceres hasta los 1399 en el sexo masculino y hasta los 1320 para el sexo femenino.
5. La DESVIACIÓN ESTANDAR más amplia se observa en el sexo femenino siendo de 252.2 y en el sexo masculino de 221.1, documentándose en ambas una gran variación en los datos.
6. La VARIANZA se observa que es menor en el sexo masculino que en el femenino, presentando así una menor dispersión de los datos.



Gráfico de dispersión.

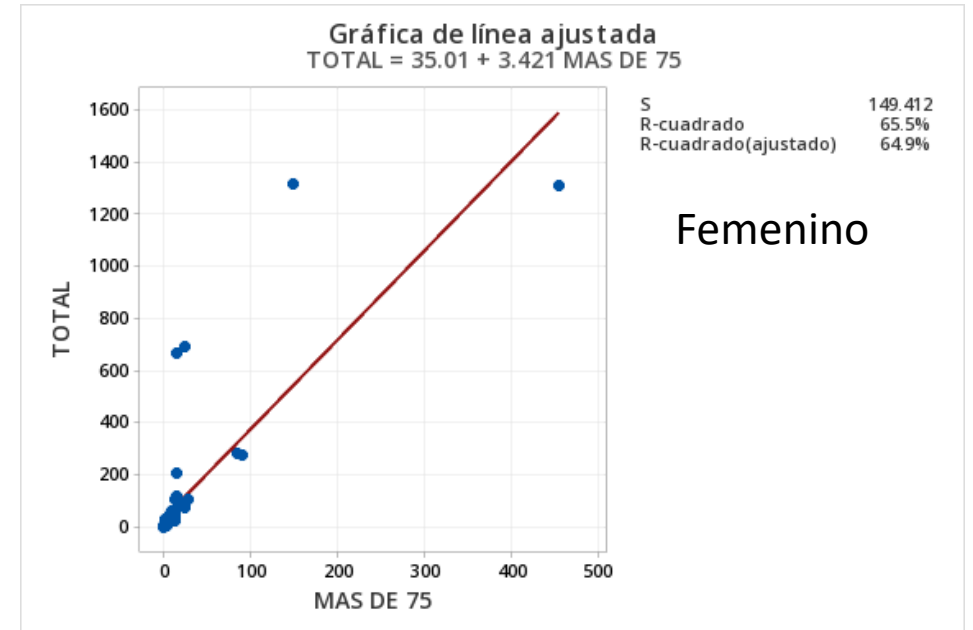
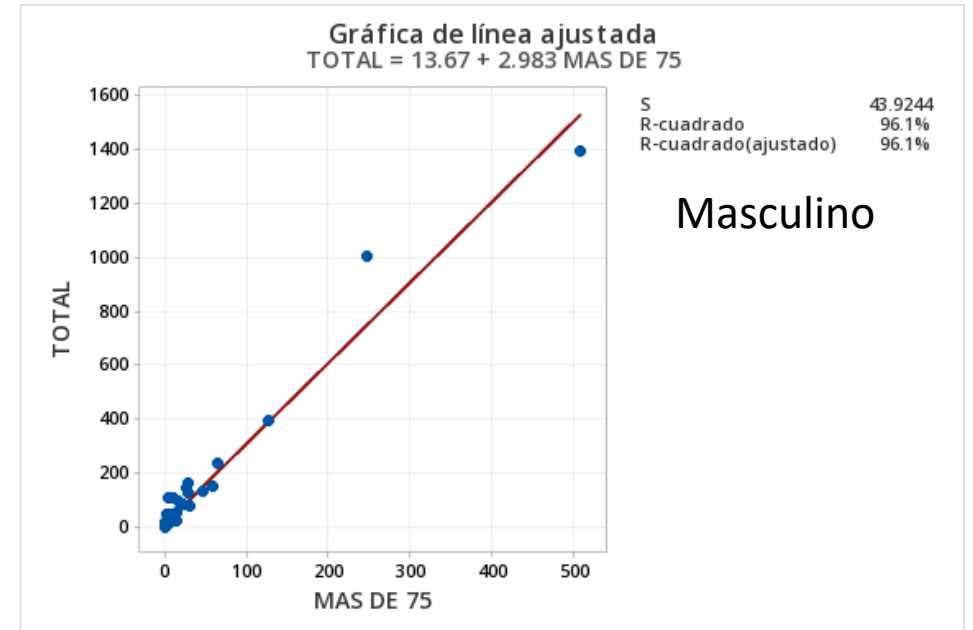


Se observa para ambos una R simple que se encuentra muy cercana a 1, por lo tanto existe una relación estrecha entre la incidencia total de cáncer y la edad de 75 años o más.



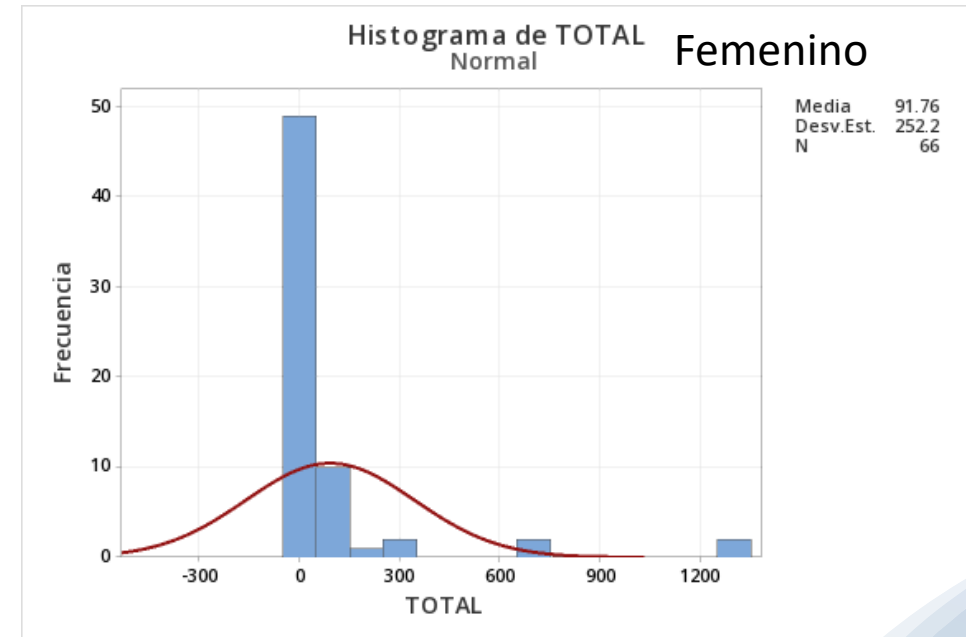
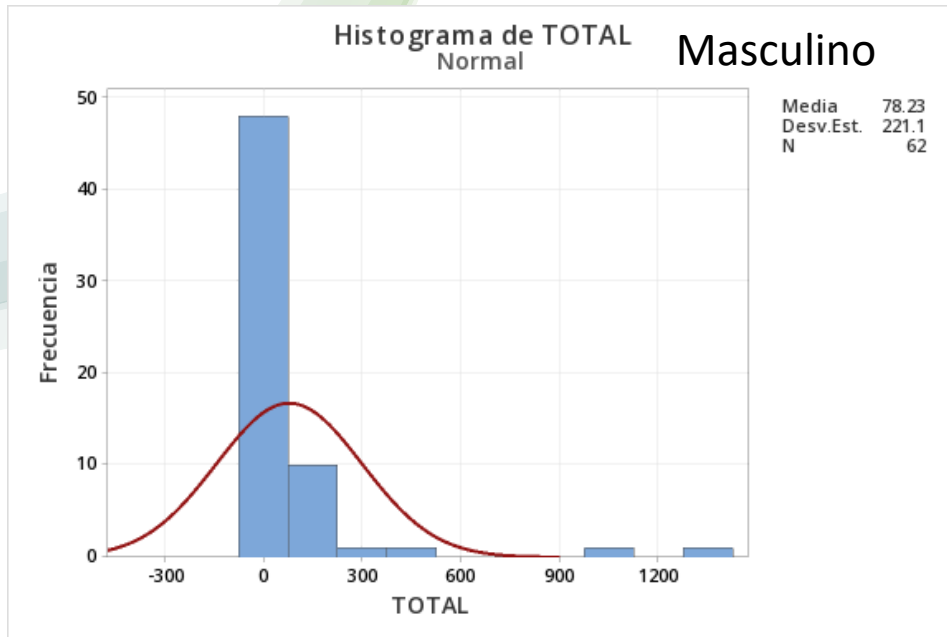
Gráfico de dispersión con línea ajustada.

1. El R cuadrado para el sexo masculino evidencia un ajuste adecuado de los datos de incidencia total de cáncer y la edad de 75 años o más al encontrarse muy cercana del 100%; mientras que para el sexo femenino lo hace de forma parcial.
2. El resultado de S para el sexo femenino al ser tan elevado confirma que el modelo tiene un ajuste bajo.



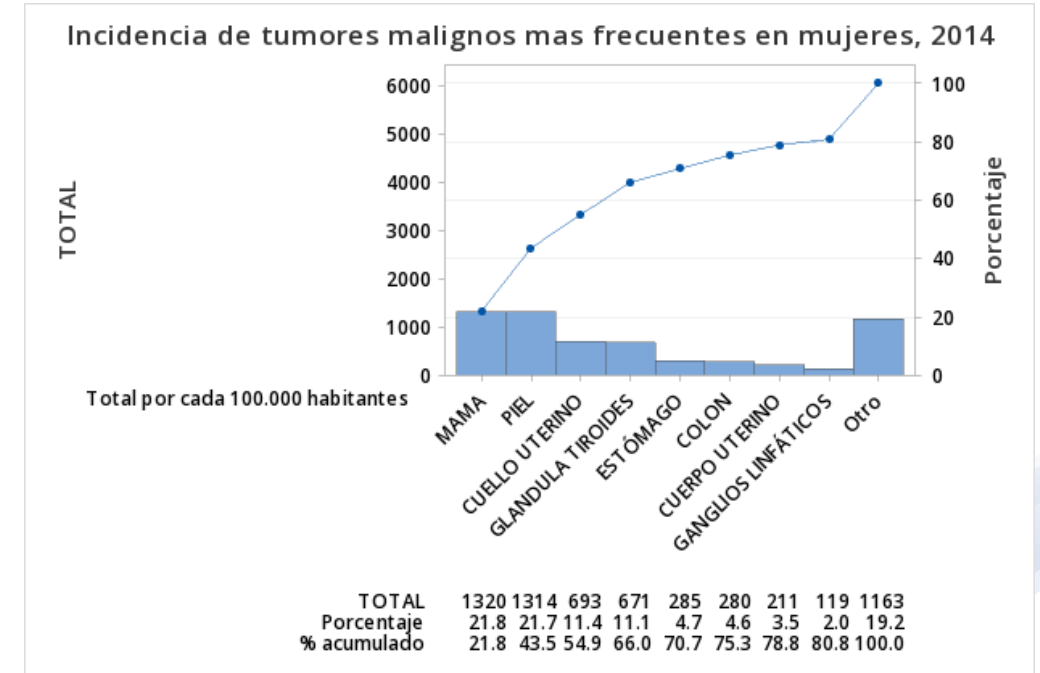
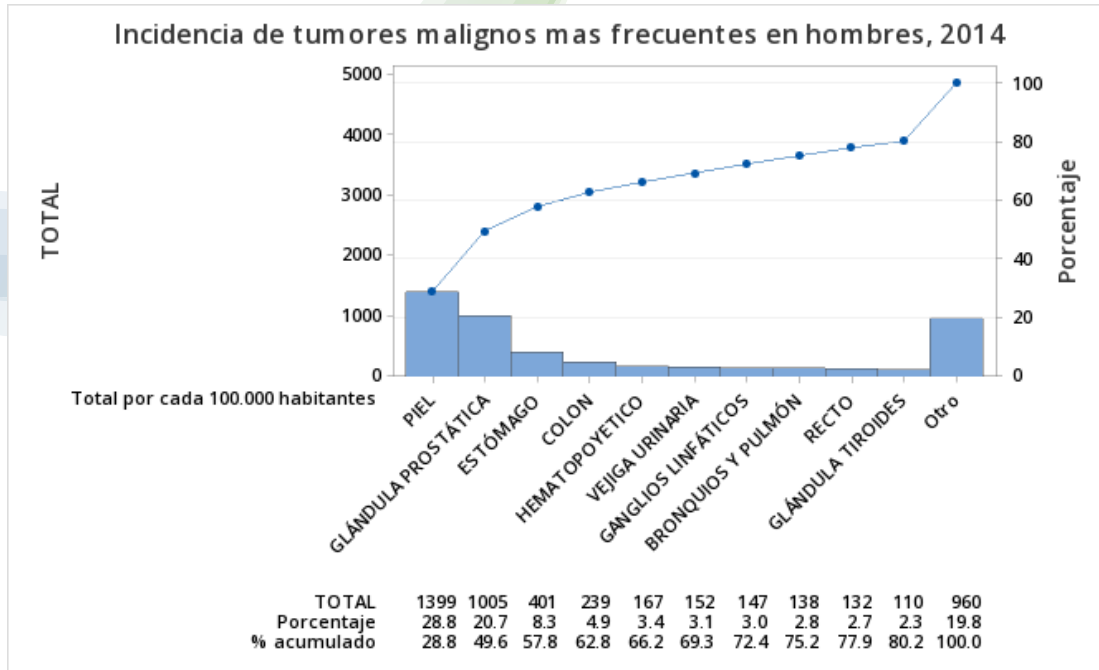


Histograma



1. Se observa que para ambos sexos hay una asimetría hacia la derecha y la presencia de barras aisladas indicando la presencia de algunos valores atípicos.
2. En estos casos los valores atípicos corresponden a una incidencia muy alta de cáncer de piel en ambos sexos y de mama en sexo femenino y de próstata en el masculino.

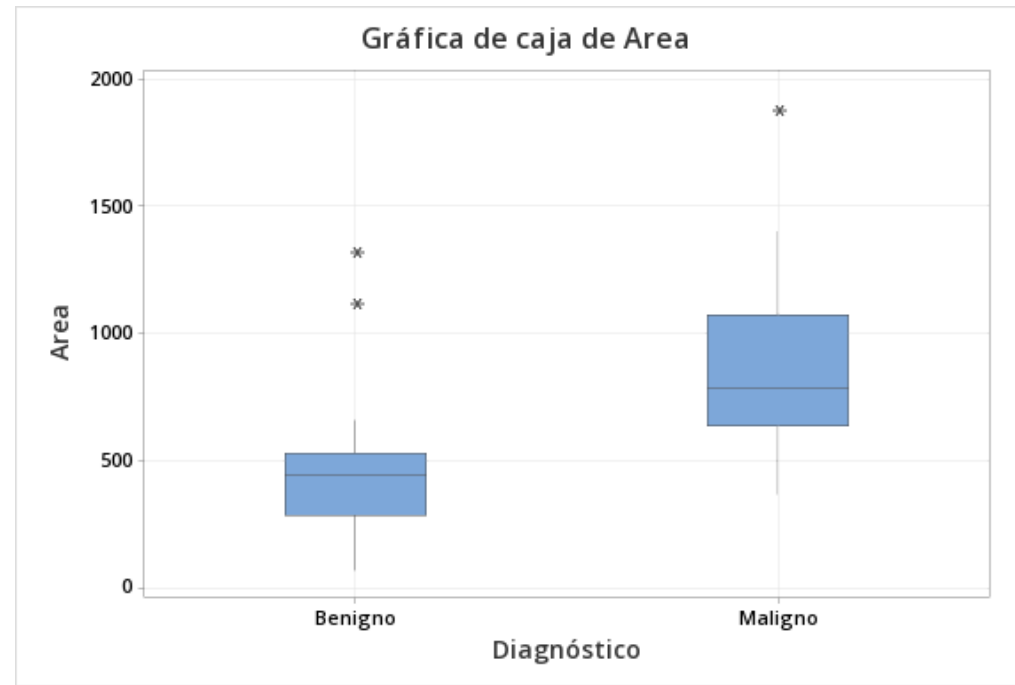
Diagrama de Pareto.



1. Se documenta que para el sexo masculino 10 tipos de cánceres representan el 80.2% del gran total y que para el sexo femenino 8 tipos de cánceres representan el 80.8% del gran total.
2. Para el sexo masculino esos 10 tipos representan la mayor incidencia, los cuales equivalen a un 16.2% de 62 tipos de cánceres.
3. Para el sexo femenino esos 8 tipos equivalen a un 12.2% de 66 tipos de cánceres.



Caja y bigotes para cáncer de próstata.

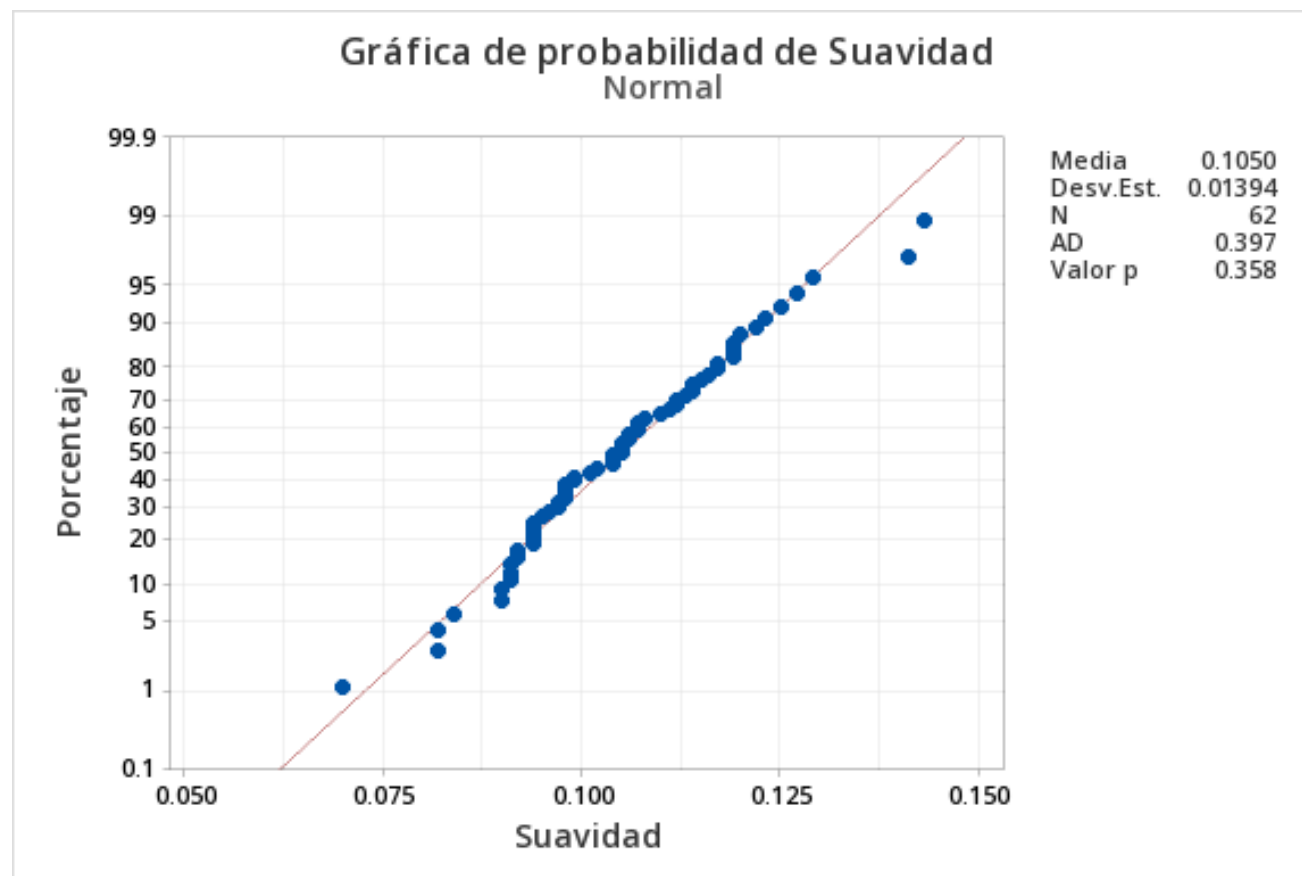


1. Se observa una mediana de 450 para los pacientes con diagnóstico de benignidad, mientras que para el diagnóstico de malignidad una mediana mayor, siendo esta de 790.
2. Para el diagnóstico de benignidad se observa un rango intercuartil de 243 y de 435 para malignidad, presentando este último una mayor dispersión y variabilidad de los datos.
3. Se observan valores atípicos en ambos diagnósticos y sería en este caso de gran importancia analizar estos en el caso del diagnóstico benigno.



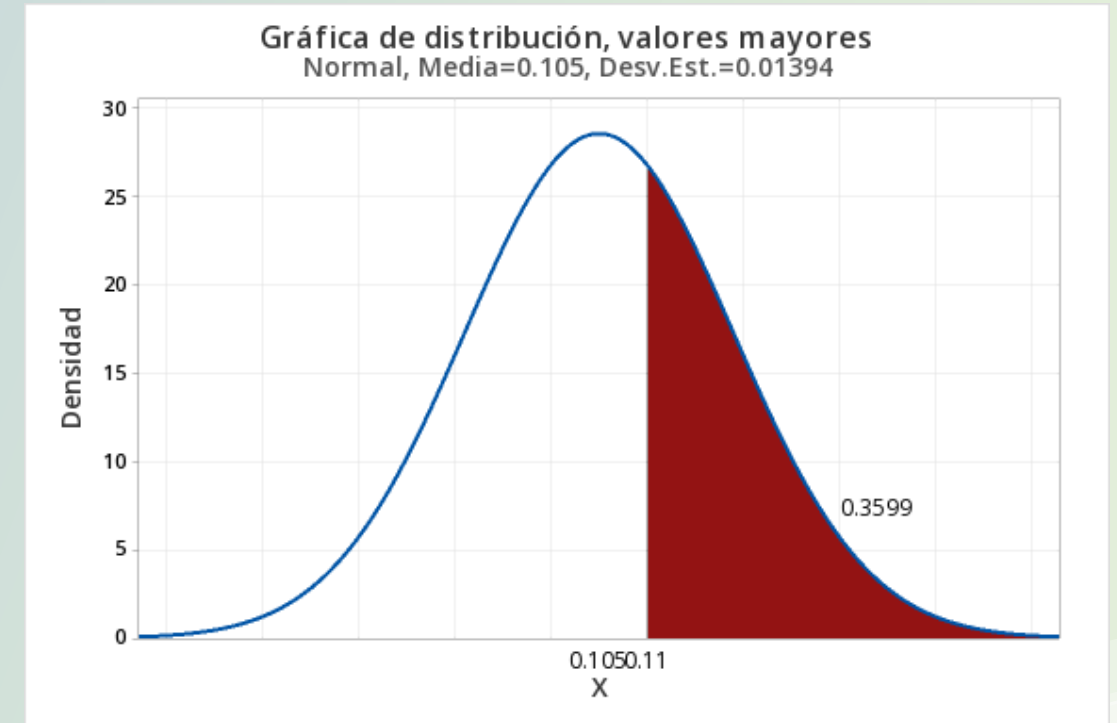
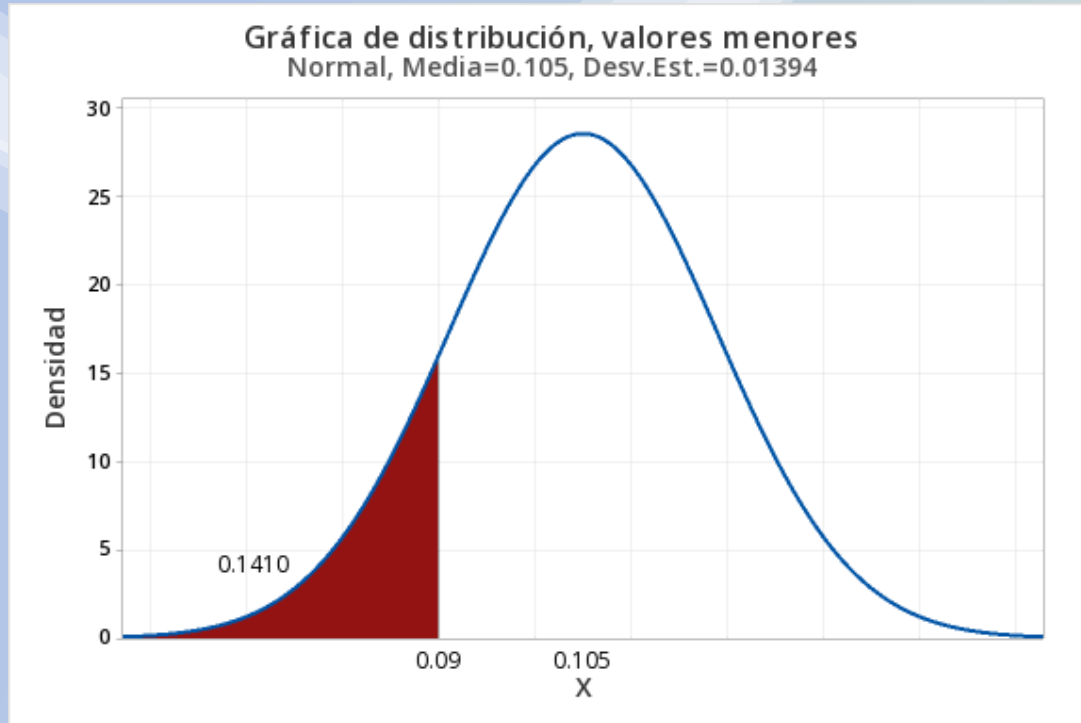
Prueba de normalidad
con distribución
normal / Base de
datos de cáncer de
próstata.

Se documenta para la
suavidad un valor de
 p con una
distribución normal.





Estudio de probabilidades.

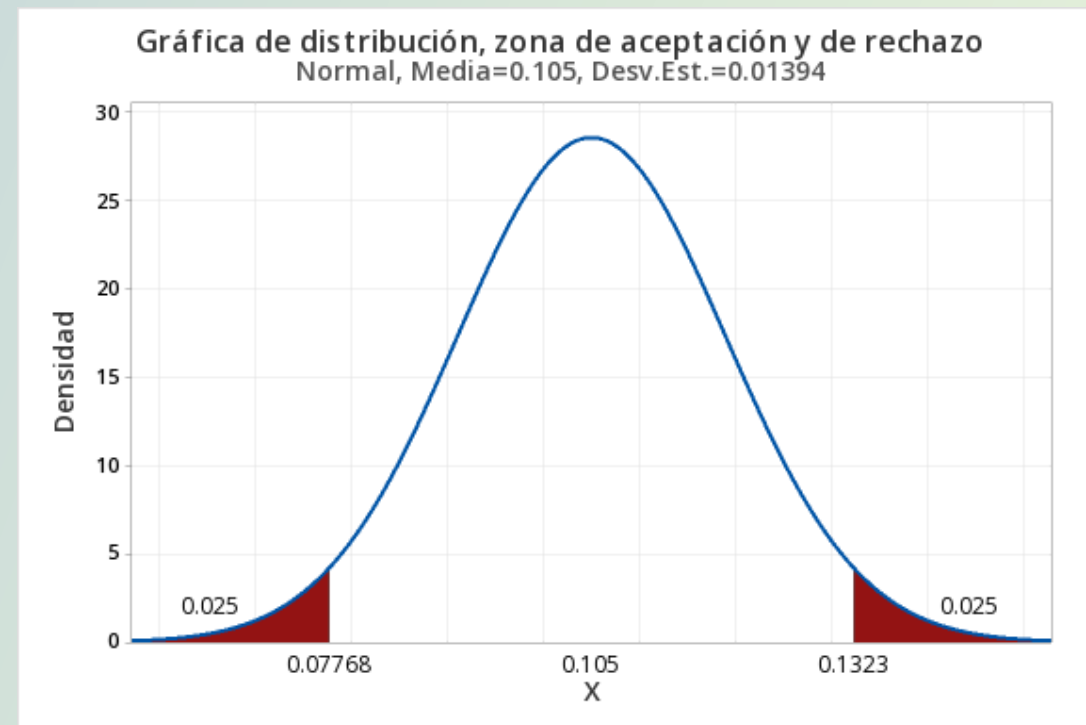
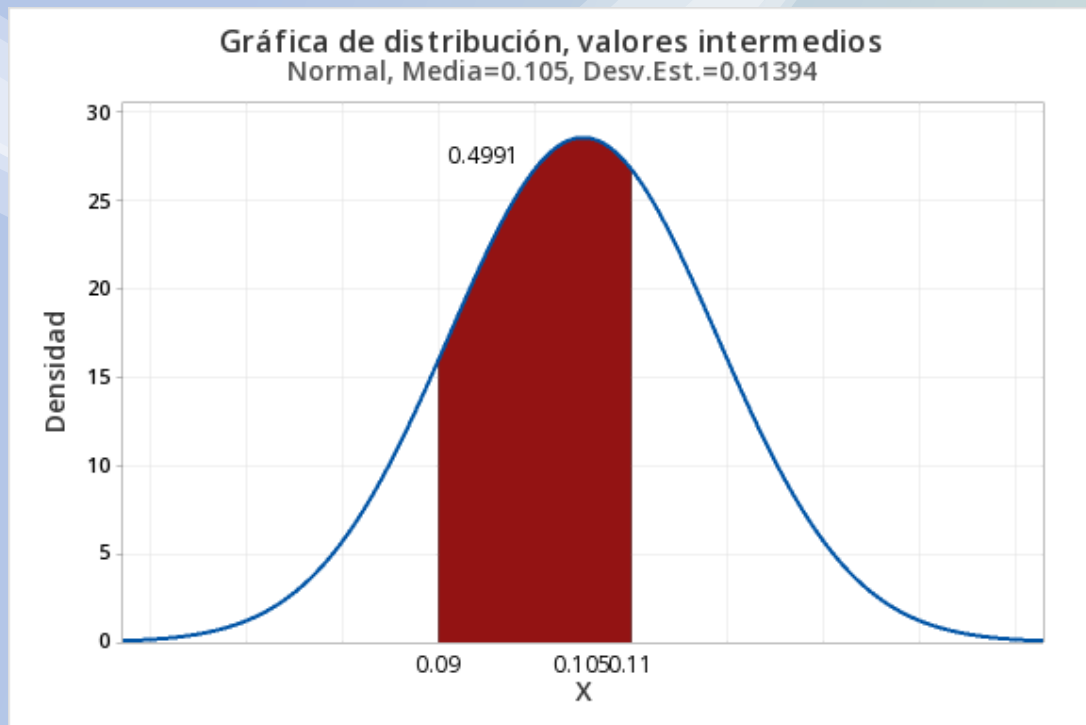


- Se puede concluir que hay una probabilidad del 14.10% de que la suavidad sea menor que 0.09, en los pacientes con diagnóstico de cáncer de próstata.

- Se puede concluir que hay una probabilidad del 35.99% de que la suavidad sea mayor que 0.11, en los pacientes con diagnóstico de cáncer de próstata.



Estudio de probabilidades.



- Se puede concluir que hay una probabilidad del 49.91% de que la suavidad esté entre 0.09 y 0.11 en los pacientes con diagnóstico de cáncer de próstata.
- Podemos concluir que encontraremos en un 95% de confianza suavidad comprendida entre 0.077 y 0.132; y tendremos en un 2.5% una suavidad menor de 0.077 y en un 2.5% mayor de 0.132.



Cálculo del tamaño de la muestra.

- 1er caso: En la evaluación de pacientes con cáncer de mama de la clínica de oncología se requiere un análisis del tamaño de la masa mamaria. Se realizó un muestreo piloto de 50 pacientes y documentó un comportamiento de los datos con distribución normal y se conocen los siguientes aspectos:
 1. El nivel de confianza que se requiere es de un 95%.
 2. El error tolerable en la medición es de 3 mm.
 3. La desviación estándar muestral obtenida de la prueba piloto es de 5 mm.
- 2ndo caso: El servicio de consulta externa de oncología está planteando un programa preventivo contra el cáncer de próstata y desea estimar la proporción de participantes y requiere estimar el tamaño muestral, considerando un error del 5% y un 95% de confianza.



Calculo del tamaño de la muestra con estimación de la media poblacional (1er caso).

Método

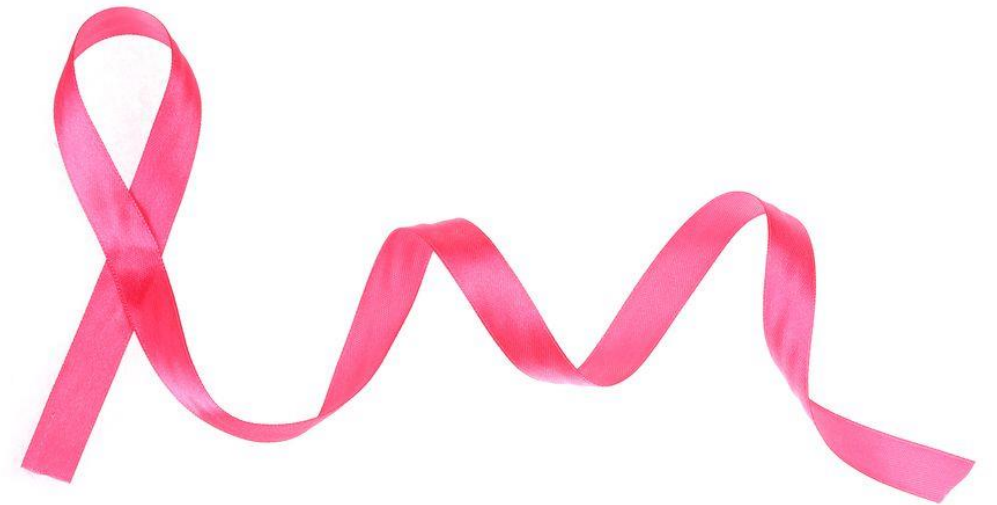
Parámetro	Media
Distribución	Normal
Desviación estándar	5 (valor de población)
Nivel de confianza	95%
Intervalo de confianza	Bilateral

**Margen
de error**

3

**Tamaño
de la
muestra**

11





Calculo del tamaño de la muestra con estimación de la proporción poblacional (2ndo caso).

Método

Parámetro	Media
Distribución	Normal
Desviación estándar	0.5 (valor de población)
Nivel de confianza	95%
Intervalo de confianza	Bilateral

**Margen
de error**

0.05

**Tamaño
de la
muestra**

385



GRACIAS