Material de Repaso para Examen Diagnóstico en Estadística

1 Medidas Descriptivas

- Media, mediana, y moda
- Rango, varianza, y desviación estándar
- Coeficiente de variación
- Asimetría y curtosis

1.1 Media, Mediana y Moda

Ejercicios:
Datos: 10, 15, 10, 20, 15, 25, 30.
Solución:

- **Media:** $\frac{10+15+10+20+15+25+30}{7} = \frac{125}{7} \approx 17.86$
- **Mediana:** Ordenamos los datos (10, 10, 15, 15, 20, 25, 30). El dato del medio es 15.
- **Moda:** El número que más se repite es 10 y 15 (ambos se repiten dos veces), así que son bimodales.

1.2 Rango, Varianza y Desviación Estándar

Datos: 8, 12, 16, 24, 28.
Solución:

- **Rango:** $28 - 8 = 20$.
- **Varianza:**
  - Media $= \frac{8+12+16+24+28}{5} = \frac{88}{5} = 17.6$.
  - Varianza $= \frac{(8-17.6)^2+(12-17.6)^2+(16-17.6)^2+(24-17.6)^2+(28-17.6)^2}{5} = 74.4$.
- **Desviación Estándar:** $\sqrt{74.4} \approx 8.63$.

1.3 Coeficiente de Variación

Datos: 7, 9, 12, 15, 18.

Solución:

- **Media:** $\frac{7+9+12+15+18}{5} = \frac{61}{5} = 12.2$.
- **Desviación Estándar:** Supongamos que es D.
- **Coeficiente de Variación:** $\frac{D}{12.2} \times 100\%$.

1.4 Asimetría y Curtosis

Datos: 1, 2, 2, 3, 4, 4, 4, 5, 6.

Solución:

- **Asimetría:** Se calcula utilizando la fórmula de asimetría.
- **Curtosis:** Se calcula utilizando la fórmula de curtosis.

2 Propiedades de los Estimadores Puntuales

- Sesgo (bias) y consistencia
- Eficiencia
- Suficiencia

2.1 Sesgo (Bias) y Consistencia

Ejercicio: Sea $\hat{\theta}$ un estimador de un parámetro θ. Demuestre que si $E[\hat{\theta}] = \theta$ entonces $\hat{\theta}$ es insesgado. Además, explique qué significa que un estimador sea consistente.

Solución:

- Un estimador $\hat{\theta}$ es insesgado si su valor esperado es igual al parámetro real θ. Es decir, $E[\hat{\theta}] = \theta$.
- Un estimador es consistente si, a medida que el tamaño de la muestra aumenta, la probabilidad de que el estimador se aproxime al valor real del parámetro tiende a uno.

2.2 Eficiencia

Ejercicio: Explique qué es un estimador eficiente y cómo se relaciona con la varianza de otros estimadores insesgados.

Solución:

- Un estimador es eficiente si tiene la menor varianza posible entre todos los estimadores insesgados para un tamaño de muestra dado.
- Esto significa que, entre todos los estimadores insesgados, el estimador eficiente proporciona la estimación más precisa del parámetro.

2.3 Suficiencia

Ejercicio: Defina qué es un estimador suficiente y proporcione un ejemplo de cómo se determina si un estimador es suficiente.
Solución:

- Un estimador es suficiente si no se puede obtener más información sobre el parámetro a estimar de la muestra que la que proporciona el estimador.
- Por ejemplo, en el caso de una distribución normal con varianza conocida, la media muestral es un estimador suficiente para la media poblacional.

3 Intervalos de Confianza

- Intervalos de confianza para medias y proporciones
- Niveles de confianza y significancia
- Interpretación de intervalos de confianza

3.1 Intervalos de Confianza para Medias y Proporciones

Ejercicio 1: Suponga que se tiene una muestra de tamaño 40 de una población normalmente distribuida con una desviación estándar conocida de 5. La media de la muestra es 50. Calcule el intervalo de confianza del 95% para la media poblacional.
Solución:

- Utilizando la fórmula del intervalo de confianza para la media con desviación estándar conocida y un nivel de confianza del 95%, obtenemos el intervalo.

Ejercicio 2: Calcule el intervalo de confianza del 90% para una proporción, dada una muestra de tamaño 200 con 120 éxitos.
Solución:

- Utilizamos la fórmula del intervalo de confianza para proporciones y un nivel de confianza del 90% para encontrar el intervalo.

3.2 Niveles de Confianza y Significancia

Ejercicio: Explique la diferencia entre el nivel de confianza y el nivel de significancia en el contexto de los intervalos de confianza.

Solución:

- El nivel de confianza indica la proporción de veces que el intervalo de confianza contiene el parámetro real cuando el experimento se repite un número infinito de veces.
- El nivel de significancia, denotado comúnmente como α, es el complemento del nivel de confianza y representa la probabilidad de cometer un error de Tipo I (rechazar la hipótesis nula cuando es verdadera).

3.3 Interpretación de Intervalos de Confianza

Ejercicio: Explique cómo se debe interpretar un intervalo de confianza del 95% para la media poblacional.

Solución:

- Un intervalo de confianza del 95% significa que si tomamos muchas muestras y calculamos el intervalo de confianza para cada una, esperamos que aproximadamente el 95% de estos intervalos contenga la media poblacional real.
- No significa que haya un 95% de probabilidad de que la media poblacional real se encuentre dentro del intervalo calculado a partir de una sola muestra.

4 Prueba de Hipótesis

- Hipótesis nula y alternativa
- Errores Tipo I y Tipo II
- Valor P
- Pruebas t y z

4.1 Hipótesis Nula y Alternativa

Ejercicio 1: Dado un estudio sobre la eficacia de un nuevo medicamento, formule la hipótesis nula y la hipótesis alternativa.

Solución:

- **Hipótesis Nula (H0):** El medicamento no tiene efecto, o el efecto es igual al del tratamiento actual.
- **Hipótesis Alternativa (H1):** El medicamento tiene un efecto significativamente diferente (mejor o peor) que el tratamiento actual.

4.2 Errores Tipo I y Tipo II

Ejercicio 2: Explique la diferencia entre un error Tipo I y un error Tipo II en el contexto de pruebas de hipótesis.

Solución:

- **Error Tipo I:** Ocurre cuando rechazamos la hipótesis nula siendo esta verdadera (falso positivo).
- **Error Tipo II:** Ocurre cuando no rechazamos la hipótesis nula siendo esta falsa (falso negativo).

4.3 Valor P

Ejercicio 3: Defina qué es el valor P en una prueba de hipótesis y cómo se interpreta.

Solución:

- El valor P es la probabilidad de obtener un resultado al menos tan extremo como el observado, asumiendo que la hipótesis nula es cierta.
- Un valor P bajo sugiere que es improbable observar los datos actuales si la hipótesis nula es cierta, lo que podría llevar a rechazar la hipótesis nula.

4.4 Pruebas t y z

Ejercicio 4: Explique en qué situaciones se utilizaría una prueba t en lugar de una prueba z.

Solución:

- La prueba t se usa generalmente cuando el tamaño de la muestra es pequeño (n ¡ 30) y la desviación estándar poblacional no es conocida.
- La prueba z se utiliza cuando el tamaño de la muestra es grande ($n \geq 30$) o la desviación estándar poblacional es conocida.