

UNIVERSIDAD NACIONAL DEL ALTIPLANO

Escuela Profesional de Ingeniería Estadística e Informática

Comparación Entre Prueba t-Student y Análisis de Varianza (ANOVA)

Estadística Computacional

CURSO: Estadística Computacional

ESTUDIANTE: Edgar Jeferson Cusihuaman Garate

DOCENTE: Fred Torres Cruz

Puno, Perú
28 de mayo de 2025

1. Introducción

En el análisis estadístico, frecuentemente necesitamos comparar medias entre grupos para determinar si las diferencias observadas son estadísticamente significativas. Dos de las técnicas más utilizadas para este propósito son la prueba t-Student y el Análisis de Varianza (ANOVA).

Este documento presenta una comparación detallada entre ambas técnicas, analizando sus fundamentos teóricos, aplicaciones, ventajas y limitaciones, con el objetivo de proporcionar una guía práctica para la selección del método más apropiado según el contexto de investigación.

2. Prueba t-Student

2.1. Definición y Fundamentos

La prueba t-Student es una prueba estadística paramétrica utilizada para comparar medias cuando tenemos muestras pequeñas ($n \leq 30$) o cuando la desviación estándar poblacional es desconocida. Fue desarrollada por William Sealy Gosset bajo el seudónimo "Student".^{en 1908.}

2.2. Tipos de Prueba t

1. **Prueba t para una muestra:** Compara la media de una muestra con un valor hipotético.
2. **Prueba t para muestras independientes:** Compara las medias de dos grupos independientes.
3. **Prueba t para muestras pareadas:** Compara las medias de dos mediciones en los mismos sujetos.

2.3. Fórmula del Estadístico t

Para muestras independientes:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (1)$$

Donde:

- \bar{x}_1, \bar{x}_2 = medias muestrales
- s_p = desviación estándar pooled

- $n_1, n_2 =$ tamaños de muestra

2.4. Supuestos de la Prueba t

- Normalidad de la distribución
- Independencia de las observaciones
- Homogeneidad de varianzas (para muestras independientes)

3. Análisis de Varianza (ANOVA)

3.1. Definición y Fundamentos

ANOVA es una técnica estadística desarrollada por Ronald Fisher que permite comparar las medias de tres o más grupos simultáneamente. Su nombre deriva del hecho de que analiza las varianzas para hacer inferencias sobre las medias.

3.2. Tipos de ANOVA

1. **ANOVA de un factor:** Compara medias de varios grupos basándose en un solo factor.
2. **ANOVA de dos factores:** Considera dos factores y su posible interacción.
3. **ANOVA de medidas repetidas:** Para datos donde se realizan múltiples mediciones en los mismos sujetos.

3.3. Fórmula del Estadístico F

$$F = \frac{MSB}{MSW} = \frac{\frac{SSB}{k-1}}{\frac{SSW}{N-k}} \quad (2)$$

Donde:

- MSB = Media cuadrática entre grupos
- MSW = Media cuadrática dentro de grupos
- SSB = Suma de cuadrados entre grupos
- SSW = Suma de cuadrados dentro de grupos
- k = número de grupos
- N = número total de observaciones

3.4. Supuestos de ANOVA

- Normalidad de la distribución en cada grupo
- Independencia de las observaciones
- Homogeneidad de varianzas (homocedasticidad)

4. Comparación Entre t-Student y ANOVA

4.1. Tabla Comparativa

Cuadro 1: Comparación entre Prueba t-Student y ANOVA

Criterio	Prueba t-Student	ANOVA
Número de grupos	Máximo 2 grupos	3 o más grupos (también aplicable a 2)
Estadístico	Distribución t	Distribución F
Hipótesis nula	$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$
Comparaciones múltiples	No aplica	Requiere pruebas post-hoc
Error Tipo I	Controlado para una comparación	Controlado para múltiples comparaciones
Potencia estadística	Mayor para 2 grupos	Menor para 2 grupos, mayor para múltiples
Complejidad	Menor	Mayor

4.2. Ventajas y Desventajas

4.2.1. Prueba t-Student

Ventajas:

- Simplicidad en cálculo e interpretación
- Mayor potencia estadística para comparar dos grupos
- Intervalos de confianza directos para la diferencia de medias
- Robusta ante pequeñas violaciones de supuestos

Desventajas:

- Limitada a comparaciones entre dos grupos
- Aumento del error Tipo I en comparaciones múltiples
- No proporciona información sobre variabilidad total

4.2.2. ANOVA

Ventajas:

- Permite comparar múltiples grupos simultáneamente
- Controla el error Tipo I familiar
- Proporciona información sobre fuentes de variación
- Flexible para diseños experimentales complejos

Desventajas:

- Mayor complejidad computacional
- Requiere pruebas post-hoc para identificar diferencias específicas
- Menos potente que t-Student para dos grupos
- Más sensible a violaciones de supuestos

5. Criterios de Selección

5.1. Cuándo Usar Prueba t-Student

- Comparación entre exactamente dos grupos
- Cuando se requiere máxima potencia estadística
- Análisis exploratorios o confirmatorios simples
- Cuando se necesita un intervalo de confianza para la diferencia

5.2. Cuándo Usar ANOVA

- Comparación entre tres o más grupos
- Cuando se quiere controlar el error Tipo I familiar
- Diseños experimentales con múltiples factores
- Cuando se requiere analizar fuentes de variación

6. Ejemplo Práctico

Consideremos el estudio sobre niveles de energía y consumo de café presentado anteriormente, donde se compararon dos grupos:

- Grupo con café: $n = 14$, $\bar{x} = 8,14$, $s = 0.77$
- Grupo sin café: $n = 15$, $\bar{x} = 4,73$, $s = 0.88$

6.1. Aplicación de Prueba t-Student

El resultado obtenido fue:

- $t = 11,094$
- Grados de libertad = 26.885
- Valor $p = 1,543 \times 10^{-11}$
- IC 95 % = [2.778, 4.040]

6.2. Si Aplicáramos ANOVA

Para los mismos datos, ANOVA proporcionaría:

- $F = t^2 = (11,094)^2 = 123,08$
- Grados de libertad = (1, 27)
- Valor p idéntico = $1,543 \times 10^{-11}$

Nota importante: Para dos grupos, $F = t^2$ y ambas pruebas son equivalentes.

7. Consideraciones Prácticas

7.1. Software Estadístico

Ambas pruebas están ampliamente implementadas en software estadístico:

- R: `t.test()`, `aov()`, `anova()`
- Python: `scipy.stats.ttest_ind()`, `scipy.stats.f_oneway()`
- SPSS: Analyze → Compare Means, Analyze → General Linear Model
- SAS: PROC TTEST, PROC ANOVA

7.2. Verificación de Supuestos

Es crucial verificar los supuestos antes de aplicar cualquiera de las pruebas:

- Normalidad: Prueba de Shapiro-Wilk, gráficos Q-Q
- Homogeneidad de varianzas: Prueba de Levene, Bartlett
- Independencia: Análisis de residuos, diseño experimental

8. Conclusiones

La elección entre prueba t-Student y ANOVA depende principalmente del número de grupos a comparar y los objetivos específicos del análisis:

1. Para comparaciones entre dos grupos, la prueba t-Student es más directa y potente.
2. Para tres o más grupos, ANOVA es la opción adecuada para controlar el error Tipo I.
3. Ambas pruebas requieren la verificación cuidadosa de sus supuestos.
4. La interpretación debe considerar tanto la significancia estadística como la relevancia práctica.

En el contexto del estudio sobre consumo de café y niveles de energía, la prueba t-Student fue la elección apropiada dado que se compararon únicamente dos grupos, proporcionando evidencia clara de una diferencia estadísticamente significativa y prácticamente relevante.