Introducción a las técnicas heurísticas

Dr. Roman Anselmo Mora Gutierréz

5 de enero de 2020

Universidad Autónoma Metropolitana

Tabla de contenido

Introducción

Ideas básicas

Probablemente, ya tengan alguna idea de lo que es: optimización, un problema, una instancia, etc. Sin embargo, es necesario revisar y analizar dichas ideas, pues serán utilizadas de manera recurrente a lo largo de este curso.

Optimización

La idea de optimización es fundamental en diversas disciplinas del conocimiento, tales como: investigación de operaciones, administración, finanzas, telecomunicaciones, etc. En otras palabras, en toda aquella rama del conocimiento donde un tomador de decisiones deba elegir, dentro de un conjunto de opciones, la mejor solución de acuerdo a criterios preestablecidos.

A continuación, se exponen algunas definiciones de optimización.

1 Luenberger precisa que la optimización es uno de los principios básicos del análisis de problemas[1] complejos de decisión y su proceso consiste en la asignación de valores a un conjunto de variables interrelacionadas, centrando la atención en un mecanismo diseñado para cuantificar la calidad de la decisión [1].

---

[1] Un problema es una diferencia, desviación o un desequilibrio entre el estado real e ideal de un sistema, además de ser lo suficientemente importante para justificar su resolución.

Optimización

2 Hall expresa que la optimización es lograr la mejor armonía entre el sistema y sus integrantes y su proceso comprende desde el planteamiento de un problema hasta el análisis y selección de la mejor alternativa [2].

3 La optimización es seleccionar de un conjunto de alternativas posibles a la mejor de ellas con base en algún criterio de decisión [3].

4 Optimización es obtener la mejor solución posible de una actividad o un proceso a través del uso adecuado de la información y conocimientos disponibles.

Optimización

5 La optimización (también denominada “programación matemática”) es una parte de la investigación de operaciones[2] que trata de resolver problemas de decisión en los que se deben determinar las acciones que optimicen un determinado objetivo, pero satisfaciendo ciertas limitaciones en los recursos disponibles [6].

6 La “programación matemática” es una potente técnica de modelado usada en el proceso de toma de decisiones [7].

[2] La investigación de operaciones es una rama de las matemáticas aplicadas, que consistente en el uso del enfoque científico en la toma de decisiones con el objeto de mejorar el diseño u operación de un sistema [4]. Ackoff la define como la aplicación por grupos interdisciplinarios del método científico a problemas relacionados con el control de las organizaciones o sistemas a fin de que se produzcan soluciones que mejor sirvan a los objetivos de toda la organización [5].

Optimización

De manera general, se puede definir a la optimización como el conjunto de conocimientos, principios, teorías, técnicas y herramientas útiles y necesarias para resolver problemas de programación matemática.

Se debe mencionar, que la resolución de un problema es un proceso racional que involucra desde identificar el problema de interés hasta la elección y ejecución de un conjunto de acciones a fin de eliminarlo o reducirlo. Este proceso debe ser sistemático y guiado por el conocimiento disponible sobre el sistema.

¿ Qué problemas se pueden optimizar?

Los problemas que se pueden optimizar de manera general contienen los siguientes elementos:

- **Alternativas o variables de decisión**: Son n decisiones cuantificables, cuyo valor afecta el desempeño del sistema.
- **Restricciones**: Representan un conjunto de m relaciones o condiciones (expresadas como ecuaciones e desigualdades) que deben satisfacer un subconjunto de variables.
- **Función objetivo (o funciones objetivo)**: Es una medida cuantitativa sobre la calidad de las soluciones de un problema. Se expresa como una función matemática de las variables de decisión.

¿ Qué problemas se pueden optimizar?

Por ende y sin perdida de generalidad, un "problema de optimización" se puede modelar como:

$$\text{máx}\, f(x)$$

$$\text{sujeto a:}$$

$$x \in F \subseteq S$$

$$S \subseteq \mathbb{R}^n$$

donde: x denota el conjunto de las n variables de decisión; S es el espacio de búsqueda, el cual corresponde al conjunto de soluciones o configuraciones posibles [8]; $f(x)$ es el conjunto (no vacío) de funciones objetivo, las cuales servirán como un instrumento para evaluar a cada solución candidata; F es el conjunto de soluciones factibles, el cual es un subconjunto de S cuyos elementos satisfacen una serie de restricciones impuestas por el problema [8].

Espacio de búsqueda versus espacio factible

En los problemas de optimización se definen de manera implícita el espacio de búsqueda $\mathcal{S}$ al definir las variables de decisión. $\mathcal{S}$ es un conjunto finito o infinito de puntos, integrado por todas las soluciones candidatas; es decir, este conjunto puede verse como un rectángulo $n-$dimensional generado por la intersección del dominio de cada una de las variables [9]. En contraste, el espacio factible $\mathcal{F}$ es $\mathcal{F} \subseteq \mathcal{S}$ definido por un conjunto adicional de m restricciones ($m \geq 0$).

Espacio de búsqueda versus espacio factible

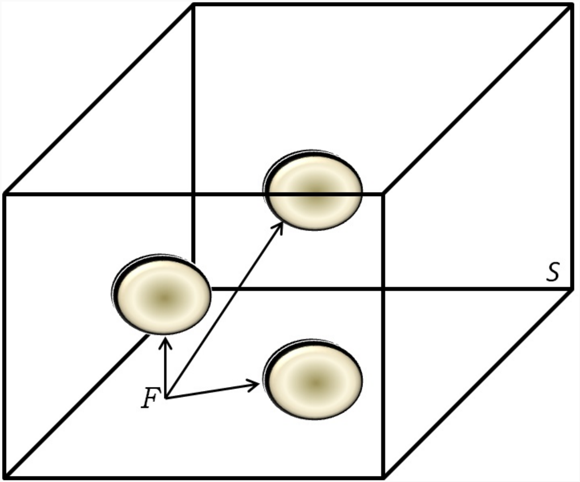


Figura 1: Espacio de búsqueda $\mathcal{S}$ y región factible $\mathcal{F}$ Fuente [10].

¿ Qué es modelar?

En la definición anterior se usó el concepto “modelar”, por lo cual es importante precisar este concepto.

Modelar es un proceso creativo-intelectual para la generación de modelos, el cual debe ser sistemático, racional y teóricamente guiado, su objetivo es analizar y resolver problemas. La modelación de problemas de optimización ha sido abordada, estudiada y sitematizada por la investigación de operaciones.

¿ Qué es modelar?

Un “modelo” es una representación o abstracción selectiva (cuantitativa o cualitativa) de las características de un sistema. Todo problema de optimización debe ser formulado a través de un “modelo matemático”; ya que estos modelos describen de modo conciso y sin ambigüedad las relaciones o condiciones del problema a resolver por medio del lenguaje y estructuras matemáticas; lo cual permite emplear técnicas matemáticas y computacionales de alto poder, para analizar y resolver dicho problema.

Definiciones adicionales

Definición

Un problema de optimización combinatoria está dado por un conjunto de instancias del problema y es un problema de maximización o un problema de minimización [11].

Definiciones adicionales

Definición

Una instancia de un problema de optimización combinatoria es una pareja (F, f)*, donde el conjunto de soluciones* F *es el conjunto de soluciones factibles y* $f : F \to \Re$ *es la función objetivo. El problema consiste en encontrar una solución óptima global,* $i.e.$*, una* $i^* \in F$ *tal que* $f(i^*) \leq f(i)$ *para toda* $i \in F$*. Más aún,* $f^* = f(i^*)$ *denota el costo óptimo y* $F^* = \{i \in F | f(i) = f^*\}$ *denota el conjunto de soluciones óptimas [11].*

Ver ejemplo de programación lineal....

Algoritmos

Definición
Un algoritmo es un procedimiento computacional bien definido, que toma algunos valores o conjunto de valores como entrada y produce algún valor o conjunto de valores, como salida

Preguntas a contestar

- ¿Cómo medir la eficiencia de un algoritmo?
- ¿Cómo medir la dificultad de un problema?
- ¿Cómo saber si un algoritmo es óptimo para un problema? (*i.e.*, ¿cómo es posible saber que no existe otro algoritmo más eficiente para resolver el mismo problema?)

Eficiencia

- Significa conocer los recursos que requiere el algoritmo
  - Frecuentemente, se quiere medir el tiempo de cómputo
  - Ocasionalmente, espacio en memoria, ancho de banda de comunicación
- El tiempo de cómputo o la memoria utilizada son recursos acotados

Importancia de la eficiencia

- La eficiencia de un algoritmo es equivalente o más importante que la potencia de la computadora usada
  - Un buen algoritmo implementado en una máquina lenta puede ejecutarse de manera más eficiente que un mal algoritmo implementado en una máquina rápida
  - Ejemplo: *quicksort* implementado en una *Intel 486* más eficiente que ordenamiento por inserción en una supercomputadora *IBM SP2* a partir de un tamaño de instancia mayor a 400

Tiempo de cómputo

- Medida usual de eficiencia: velocidad o tiempo de cómputo
- Medición empírica del tiempo de cómputo no es buena opción (lenguaje, SO, compilador, habilidad del programador)
- Analizar un algoritmo, aún simple, puede ser un reto
  - combinatoria discreta
  - teoría elemental de probabilidad
  - destreza algebraica, etc.

Metodología

Para cualquier algoritmo

- Elegir un tipo particular de operación (instrucción del alg.)
- Análisis matemático para determinar el No. de veces que ocurre la operación para completar el algoritmo

Costo computacional y tamaño

- Costo de ejecución (o computacional) depende del tamaño de la instancia (n)
  - Típicamente: costo $\nearrow$ cuando $n \nearrow$
- Se determina el orden o tasa de crecimiento del tiempo que toma un algoritmo
  - Algunos órdenes de crecimiento dominan a otros
  - Por ejemplo: n^3 domina a n para $n \geq n_0 = 1$. Por lo tanto, si se requieren $n^3 + n$ pasos para ejecutar un alg., se dirá que su complejidad es del orden n^3.

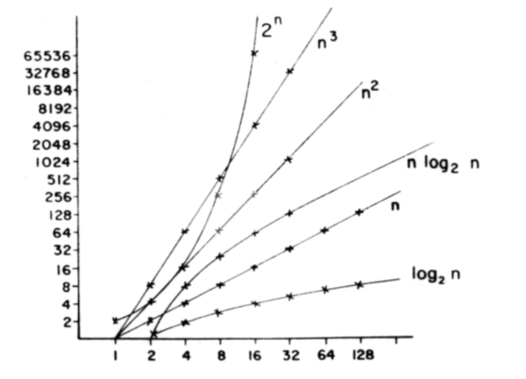
Definición

$f(n) \in \mathcal{O}(g(n))$ sii existen dos constantes positivas c y n_0 tales que $|f(n)| \leq c|g(n)|$ para toda $n \geq n_0$

- Implicaciones
  - $f(n)$ en cierto sentido está acotada por $g(n)$ cuando n es suficientemente grande
  - Si un algoritmo tiene complejidad $\mathcal{O}(g(n))$, se requiere ejecutar menos de $c|g(n)|$ veces la instrucción primaria
  - Importancia de la constante n_0
    - Ejemplo: comparar las tasas de crecimiento de $0.1n^2$ y $0.1n^2 + n + 100$ ($n = \{10, 20, 50, 100, 1000\}$)

Complejidad de los algoritmos

- Algunas clases importantes de orden de crecimiento



Problema de ordenamiento

- Entrada: una secuencia de n números $< a_1, a_2, \ldots, a_n >$
- Salida: Una permutación $< a'_1, a'_2, \ldots, a'_n >$ de la secuencia de entrada tal que $a'_1 \leq a'_2 \leq \ldots \leq a'_n$
- Gran variedad de algoritmos para resolver este problema
  - Algoritmo de la burbuja
  - Ordenamiento por inserción
  - Algoritmo *quicksort*
  - Algoritmo *mergesort*
  - ...
- Ordenamiento por inserción es un algoritmo eficiente para ordenar un número pequeño de elementos

Ordenamiento por Inserción

Input: vector $A = [a_1, a_2, \ldots, a_n]$

Output: vector $A = [a'_1, a'_2, \ldots, a'_n]$ tal que $a'_1 \leq a'_2 \leq \ldots \leq a'_n$

Para j = 2 hasta n hacer

  i = j - 1

  $Store = a_j$

  Mientras $Store < a_i$ y i > 0 hacer

    $a_{i+1} = a_i$

    i = i-1

  End

  $a_{i+1} = Store$

End

Modo operativo

- Entrada: $< 5, 2, 4, 6, 1, 3 >$
- Etapas $< 5, 2, 4, 6, 1, 3 >$

  $< 2, 5, 4, 6, 1, 3 >$

  $< 2, 4, 5, 6, 1, 3 >$

  $< 2, 4, 5, 6, 1, 3 >$

  $< 1, 2, 4, 5, 6, 3 >$

  $< 1, 2, 3, 4, 5, 6 >$ $\Rightarrow$ 5 pasos

- Entrada: $< 3, 2, 1, 4, 5, 6 >$
- Etapas $< 3, 2, 1, 4, 5, 6 >$

  $< 2, 3, 1, 3, 5, 6 >$

  $< 1, 2, 3, 4, 5, 6 >$ $\Rightarrow$ 2 pasos

Modo operativo

- El No. de operaciones primarias ejecutadas depende de los datos de entrada
- Consideremos que la operación primaria es el intercambio de datos
  - Dos veces en el ciclo principal
  - Una vez por cada paso en el ciclo interno
  - Sea $d_j =$ número de intercambios en el ciclo interno
- El número de ejecuciones de la operación primaria es

$$N_{OP} = \sum_{j=2}^{n} (2 + d_j) = 2 \cdot (n-1) + \sum_{j=2}^{n} d_j$$

Distinción por caso

- Mejor caso: la secuencia ya está ordenada
  - $\forall j \in \{2, \ldots, n\}\ d_j = 0$

$$\Rightarrow N_{OP} = 2 \cdot (n-1) = \mathcal{O}(n)$$

- Peor caso: la secuencia está completamente desordenada
  - $d_2 = 1,\ d_3 = 2,\ \ldots\ d_n = n-1$ por lo cual $\sum_{j=2}^{n} d_j = \frac{n}{2} \cdot (n-1)$

$$\Rightarrow N_{OP} = \frac{1}{2} \cdot (n-1) \cdot (n+4) = \mathcal{O}\left(n^2\right)$$

Distinción por caso

- Caso promedio: estudio más detallado
  - Cuando se considera x_j, ya se han ordenado $j-1$ elementos
  - Si x_j es el número más grande, entonces no se entra al ciclo interno y $2+d_j=2$, si x_j es el segundo número más grande, entonces se entra una vez al ciclo interno y $2+d_j=3$, ..., si x_j es el número más pequeño, entonces se entra $j-1$ veces al ciclo interno y $2+d_j=j+1$
  - La probabilidad de que x_j sea el número más grande es $1/j$, la probabilidad de que x_j sea el segundo número más grande es $1/j$, etc ...

Distinción por caso

- Caso promedio: resolución
  - Por lo anterior: $E\left(2+d_j\right)=\frac{2}{j}+\frac{3}{j}+\ldots+\frac{j+1}{j}$
    O sea: $E\left(2+d_j\right)=\frac{j+3}{2}$
  - La complejidad computacional es por lo tanto

$$\sum_{j=2}^{n} 2+d_j=\sum_{j=2}^{n} \frac{j+3}{2}=\frac{1}{4}\cdot(n-1)\cdot(n+8)=\mathcal{O}\left(n^2\right)$$

Clases de algoritmos

- Cualquier algoritmo con complejidad $\mathcal{O}(p(n))$ donde $p(n)$ es una función polinomial, es un algoritmo polinomial
- Algoritmos cuya complejidad no puede ser acotada por alguna función polinomial son algoritmos exponenciales ($\mathcal{O}(2^n)$, $\mathcal{O}(n!)$)

Análisis estadístico

En varias ocasiones es resulta más sencillo y es académicamente valioso, realizar comparaciones del desempeño de diferentes algoritmos, mediante el análisis de la calidad de las soluciones que obtienen.

- Media, mediana, desviación estándar.
- Diagramas de caja
- Prueba de Bootstrap
- Prueba de Wilcoxon

Análisis estadístico

La media muestral se define generalmente como el promedio aritmético de las observaciones de una muestra. Si las N observaciones de una muestra se denotan por $x_1, x_2, ..., x_N$, entonces la media muestral se denota como $\bar{x}$ y se calcula de la siguiente manera:

$$\bar{x} = \frac{(x_1 + x_2 + ... + x_N)}{N} = \sum_{i=1}^{N} \frac{x_i}{N} \tag{1}$$

Asimismo, se puede calcular la desviación estándar muestral, denotada por s^2, dada por:

$$s^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N - 1} \tag{2}$$

Análisis estadístico

La mediana muestral se obtiene de **ordenar** las "N" observaciones de menor a mayor incluyendo los valores repetidos.

$$\text{Mediana} = \tilde{x} = \begin{cases} \text{El dato en la posición } \frac{N+1}{2}, \text{ si "}N\text{" es impar} \\ \\ \text{El promedio de los datos en las posiciones} \\ \frac{N}{2} \text{y } \frac{N}{2}+1, \text{si "}N\text{" es par} \end{cases} \quad (3)$$

Un error que se comete con frecuencia es pensar que el resultado de la operación $\frac{N+1}{2}$, o el promedio de $\frac{N}{2}$ y $\frac{N}{2}+1$, es la mediana. Debe recordarse que los resultados de estas operaciones son la posición en que se encuentra la mediana, por lo tanto sólo sirven de guía para localizarla. En los siguientes ejemplos se ilustra esta observación.

Ejemplo
Supongan que los datos $a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9$ *están ordenados de menor a mayor, entonces la mediana se calcularía de la siguiente manera:*
Caso 1 (Impar): $\frac{N+1}{2} = \frac{9+1}{2} = 5$
Por lo tanto,
Mediana $= \tilde{x} = a_5$

Ejemplo
Supongan que los datos $a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8$ *están ordenados de menor a mayor, entonces la mediana se calcularía de la siguiente manera:*
Caso 2 (Par): $\frac{N}{2} = \frac{8}{2} = 4$ *y* $\frac{N}{2} + 1 = \frac{8}{2} + 1 = 5$
Por lo tanto,
Mediana $= \tilde{x} = \frac{a_4 + a_5}{2}$

Análisis estadístico

Cuartiles el nombre que se le da a los puntos que dividen los datos ordenados en cuatro partes iguales.
El primer cuartil q_1 divide la muestra de tal forma que el 75 % de los datos son mayores que q_1 y 25 % menores.
El segundo cuartil, q_2, es idéntico a la mediana.
El tercer cuartil q_3 divide la muestra de tal forma que el 25 % de los datos son mayores que q_3 y 75 % menores.
El **rango intercuartil** es la diferencia $q_3 – q_1$.

Análisis estadístico

Para determinar cada uno de los cuartiles se deben ordenar los datos de menor a mayor y determinar los valores ubicados en las siguientes posiciones:

$$\begin{aligned} q_1 &= \text{Dato en la posición } \tfrac{n}{4}+\tfrac{1}{4}. \\ q_2 &= \text{Dato en la posición } \tfrac{n}{2}+\tfrac{1}{2}. \\ q_3 &= \text{Dato en la posición } \tfrac{3n}{4}+\tfrac{1}{4}. \end{aligned} \tag{4}$$

Donde n es el número de observaciones.
Nuevamente se debe observar que el resultado de las operaciones indica la posición en que se encuentra cada cuartil, igual que en el caso de la mediana, pero para conocer su verdadero valor se debe recurrir a los datos dados, y en algunas ocasiones será necesario obtener el promedio de dos datos contiguos.

Ejemplo
Considera los siguientes datos: $2, 3, 4, 4, 4, 7, 9, 10, 11, 11, 11, 26$.
Aplicando las fórmulas se puede determinar la posición en la que se encuentran los cuartiles con $n = 12$.

$q_1 = \text{Dato en la posición } \frac{12}{4} + \frac{1}{4} = \text{Dato en la posición } 3.25.$

$q_2 = \text{Dato en la posición } \frac{12}{2} + \frac{1}{2} = \text{Dato en la posición } 6.5.$

$q_3 = \text{Dato en la posición } \frac{3*12}{4} + \frac{1}{4} = \text{Dato en la posición } 9.25.$

Ahora se ubican los valores de cada cuartil usando los datos ordenados.

$$2, 3, \underbrace{4, 4}_{q_1}, 4, \underbrace{7, 9}_{q_2}, 10, \underbrace{11, 11}_{q_3}, 11, 26$$

Por lo tanto, $q_1 = 4 + (4 - 4) * 0.25 = 4$, $q_2 = 7 + (9 - 7) * 0.5 = 8$ *y* $q_3 = 11 + (11 - 11) * 0.25 = 11$.

Diagramas de cajas y bigotes

Este tipo de gráfica consiste en un rectángulo, donde los lados más largos muestran la distancia entre el primer y tercer cuartil de los datos. El rectángulo está dividido por un segmento vertical que indica la posición de la mediana.

Los bigotes se tratan de un par de líneas que se extienden desde cada extremo de la caja.

El bigote bajo es una línea que va del primer cuartil al menor de los datos dentro de una distancia menor o igual a (1.5)*(rango intercuartil) a partir del primer cuartil.

El bigote superior o derecho es una línea que va del tercer cuartil al mayor dato ubicado dentro de una distancia menor o igual a (1.5)*(rango intercuartil) a partir del tercer cuartil.

Puntos atípicos, cualquier dato que no se encuentre dentro de la caja o de los bigotes es marcado e identificado individualmente mediante puntos llamados puntos atípicos.

Ejemplo

La edad de un colectivo de 30 personas se presenta en la tabla 1. Emplea esta información para realizar un diagrama de caja y bigotes.

36	25	37	24	39
39	24	29	23	41
21	50	61	65	70
36	45	31	31	20
33	24	34	40	40
24	21	30	23	22

Cuadro 1: *Edades de 30 personas, ejemplo 8.*

Diagramas de cajas y bigotes

Primero se deben calcular los cuartiles, para lo cual se ordenan los datos.
20 21 21 22 23 23 24 24 24 24 25 29 30 31 31 33 34 36 36
37 39 39 40 40 41 45 50 61 65 70.

La posición del primer cuartil es $\frac{30}{4} + \frac{1}{4} = 7.75$.

Entonces $q_1 = 24$.

La posición del segundo cuartil, que coincide con la mediana, es $\frac{N}{2} + 0.5 = 15.5$.

Entonces $\tilde{x} = q_2 = \frac{31+33}{2} = 32$.

Finalmente la posición del tercer cuartil es $\frac{3N}{4} + 0.25 = 22.75$.

Con lo cual se tiene $q_3 = 39 + (40 - 39) * (0.75) = 39.75$.

Diagramas de cajas y bigotes

Por lo tanto, los límites de la caja estarán en los puntos 24 y 39.75.
Mientras que en el punto 32 se dibujará una línea para indicar la posición de la mediana.
Ahora se calcula el rango intercuartil $= 39.5 - 24 = 15.5$.
Por lo tanto se dibujarán bigotes hasta los datos que disten menos de $(1.5)(15.5) = 23.25$ del extremo más cercano de la caja.
El bigote de abajo llegará hasta el dato con el valor 20, el dato más pequeño cuya distancia al cuartil 1 es $24–20 = 4 < 23.25$.
El bigote de arriba llegará hasta el dato con el valor 61, por ser el valor más grande que dista del cuartil 3 menos que 23.25, $61 - 39.75 = 21.25$.
Los datos con valores 65 y 70 se representarán como datos atípicos.

Bootstrap

Bootstrap es una técnica de remuestreo. Quizás su nombre provenga de su carácter de tarea imposible. Se utiliza para hacer posibles tareas que podrían parecer imposibles cuando el tamaño de las muestras es muy pequeño o cuando las distribuciones están muy sesgadas. Lo usaremos para obtener intervalos de confianza.

Figura 2: Expectativa vs realidad.

Bootstrap

Para establecer el intervalo de confianza de una media, se requiere calcular la media de múltiples muestras de una población.
De esta forma, la extracción de muestras repetidas de la población nos permite hacer descripciones e inferencias estadísticas.
Para el bootstrap las muestras analizadas se extraen de una única muestra y no de la población de la que procede. Por lo tanto, cuando hablemos de una prueba de Bootstrap consideraremos que sólo hemos extraído una muestra de la población original.

Bootstrap

Para calcular un intervalo de confianza para la media de una población se realiza el siguiente procedimiento.

Paso 1) A partir de la muestra original se extrae una muestra aleatoria con reposición.

Paso 2) Se obtiene el estadístico deseado, la media, y se utiliza como estimador de la población.

Paso 3) Se repiten los dos pasos anteriores un gran número de veces, obteniendo así un número alto de estimaciones.

Paso 4) Si se desea calcular un intervalo de confianza del $(1-\alpha)\,\%$. Se deben encontrar los percentiles $\frac{\alpha}{2}$ y $1-\frac{\alpha}{2}$.

Bootstrap

El percentil es una medida de posición usada en estadística que indica, una vez ordenados los datos de menor a mayor, el valor de la variable por debajo del cual se encuentra un porcentaje dado de observaciones en un grupo de observaciones.

Calcular $\frac{n*i}{100}$, donde n es el número de elementos de la muestra e i el percentil deseado.

Sean E y D la parte entera y decimal, respectivamente, de esta división.

Entonces el *i-ésimo* percentil, P_i, está dado por:

$$P_i = \begin{cases} elemento(E+1), \text{ si} D \neq 0 \\ \\ \frac{elemento(E)+elemento(E+1)}{2}, \text{si} D = 0 \end{cases} \tag{5}$$

Bootstrap

Ejemplo
Suponga que se ha extraído la siguiente muestra de un población: 1, 2, 4, 4, 10. Calcule un intervalo de confianza del 90 % realizando un remuestreo Boostrap.

2, 1, 10, 4, 2	4, 4, 4, 4, 1
4, 10, 10, 2, 4	1, 2, 4, 4, 2
1, 4, 1, 4, 4	4, 4, 10, 10, 2
4, 1, 1, 4, 10	4, 2, 1, 4, 4
4, 4, 1, 4, 2	4, 4, 4, 4, 4
4, 10, 10, 10, 4	4, 2, 4, 1, 1
2, 4, 4, 2, 1	4, 4, 4, 2, 4
2, 4, 1, 10, 4	10, 4, 1, 4, 4
1, 10, 2, 10, 10	4, 2, 1, 1, 2
4, 1, 10, 1, 10	10, 2, 2, 1, 1

Cuadro 2: *Muestras aleatorias sin reemplazo.*

Bootstrap

	Media		Media
2, 1, 10, 4, 2	3.8	4, 4, 4, 4, 1	3.4
4, 10, 10, 2, 4	6	1, 2, 4, 4, 2	2.6
1, 4, 1, 4, 4	2.8	4, 4, 10, 10, 2	6
4, 1, 1, 4, 10	4	4, 2, 1, 4, 4	3
4, 4, 1, 4, 2	3	4, 4, 4, 4, 4	4
4, 10, 10, 10, 4	7.6	4, 2, 4, 1, 1	2.4
2, 4, 4, 2, 1	2.6	4, 4, 4, 2, 4	3.6
2, 4, 1, 10, 4	4.2	10, 4, 1, 4, 4	4.6
1, 10, 2, 10, 10	6.6	4, 2, 1, 1, 2	2
4, 1, 10, 1, 10	5.2	10, 2, 2, 1, 1	3.2

Cuadro 3: Cálculo de medias.

Se desea calcular un intervalo de confianza del 90 %. Se deben encontrar los percentiles 5 y 95.

Se ordenan las medias de forma ascendente.

2, 2.4, 2.6, 2.6, 2.8, 3, 3, 3.2, 3.4, 3.6, 3.8, 4, 4, 4.2, 4.6, 5.2, 6, 6, 6.6, 7.6.

El percentil 5 es 2.2 y el percentil 95 es 7.1. Por lo tanto el intervalo de confianza es (2.2, 7.1).

Wilcoxon

Prueba de Wilcoxon.
La prueba de los rangos con signo de Wilcoxon es una prueba no paramétrica para comparar el rango medio de dos muestras relacionadas y determinar si existen diferencias entre ellas.
Se utiliza para comparar dos mediciones relacionadas y determinar si la diferencia entre ellas se debe al azar o no. En este último caso, la diferencia será estadísticamente significativa.
Suponga que se dispone de n pares de observaciones, denominadas (x_i, y_i), que provienen de poblaciones con medias θ_1 y θ_2. El objetivo de la prueba es comprobar si $\theta_1 = \theta_2$.

Wilcoxon

Para esta prueba se plantean las siguientes hipótesis:

$$H_0 : \theta_1 - \theta_2 = 0 \tag{6}$$

$$H_1 : \theta_1 - \theta_2 \neq 0 \tag{7}$$

Cuando el número de observaciones es suficientemente grande, al menos 30, se puede usar una distribución normal para determinar si debe rechazarse la hipótesis nula. En este caso se usará la normalización:

$$Z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \tag{8}$$

Donde:
n el el número de parejas de observaciones consideradas.
W el estadístico de prueba.
$\mu = \frac{n(n+1)}{4}$ y $\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

Wilcoxon

Procedimiento.

- Calcular las diferencias en valor absoluto entre las dos puntuaciones de cada pareja. Borrar las diferencias iguales a cero y ajustar el valor de n.
- Ordenar en forma ascendente las diferencias calculadas en el paso anterior y asignarle un rango. En caso de que encontrar varias diferencias con el mismo valor, empates, se les deberá asignar el rango promedio.
- Sumar los rangos correspondientes a las diferencias positivas (W+) y los correspondientes a las diferencias negativas (W-).
- Calcular el estadístico de prueba W = max(W+,W-).
- Calcular Z y comparar con el valor crítico para el nivel de confianza requerido. Si $|Z| \leq z_{\alpha/2}$ no se puede rechazar la hipótesis nula y por lo tanto se considera que las dos muestras tienen la misma media.

Wilcoxon

Cuando hay empates, la desviación estándar sufre una modificación:

$$Z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^{g} \frac{t_i^3 - t_i}{2}}} \tag{9}$$

Donde g el el número de grupos empatados y t_i es el tamaño del *i-ésimo* grupo empatado.

El nivel de confianza, $(1-\alpha)\,\%$, para este tipo de pruebas suele ser del 90 %, 95 % o 99 %. En estos casos, se tendrá $\alpha = 0.10$, $\alpha = 0.05$ o $\alpha = 0.01$ respectivamente. Por lo tanto, al revisar las tablas para la distribución normal estándar se obtiene:

- Para un nivel de concianza del 90 % $z_{0.10/2} = 1.65$
- Para un nivel de concianza del 95 % $z_{0.05/2} = 1.96$
- Para un nivel de concianza del 99 % $z_{0.10/2} = 2.58$

Para niveles de confianza diferentes deberán consultar las tablas de distribución normal.

Ejemplo
Considere las siguientes muestras y realice una prueba de Wilcoxon para determinar si provienen de poblaciones con la misma media.

	x	y			x	y
1	2	3.5		7	14.9	16.7
2	3.6	5.7		8	6	6
3	2.6	2.9		9	2.3	3.8
4	2.6	2.4		10	2	4
5	7.3	9.9		11	6.8	9.1
6	3.4	3.3		12	8.5	20.9

Cuadro 4: *Datos para Wilcoxon.*

Observación: Por detalles de espacio y practicidad, el número de datos es pequeño, pero realizaremos un análisis como si fueran al menos 30 parejas de datos.

Wilcoxon

Calculamos las diferencias.

	x	y	x-y	$\mid x - y \mid$
1	2	3.5	-1.5	1.5
2	3.6	5.7	-2.1	2.1
3	2.6	2.9	-0.3	0.3
4	2.6	2.4	0.2	0.2
5	7.3	9.9	-2.6	2.6
6	3.4	3.3	0.1	0.1
7	14.9	16.7	-1.8	1.8
8	6	6	0	0
9	2.3	3.8	-1.5	1.5
10	2	4	-2	2
11	6.8	9.1	-2.3	2.3
12	8.5	20.9	-12.4	12.4

Cuadro 5: Prueba de Wilcoxon.

Wilcoxon

Ordenamos las diferencias y asignamos un rango.

0.1	0.1	1
0.2	0.2	2
-0.3	0.3	3
-1.5	1.5	4.5
-1.5	1.5	4.5
-1.8	1.8	6
-2	2	7
-2.1	2.1	8
-2.3	2.3	9
-2.6	2.6	10
-12.4	12.4	11

Cuadro 6: Prueba de Wilcoxon.

$$W+ = 3 \text{ y } W- = 63.$$

Wilcoxon

De esta forma tenemos lo siguiente:

- El número de observaciones disminuyó por un empate a cero, por lo tanto $n = 11$.
- Dos observaciones empataron a 1.5, por lo tanto se hará la adecuación de la desviación estándar con $t = 2$.
- El estadístico W es igual a 63.

De esta forma,

$$Z = \frac{63 - \frac{11*(11+1)}{4}}{\sqrt{\frac{11(11+1)(2*11+1)}{24} - \frac{2^3-2}{2}}} = 3.637588565$$

Wilcoxon

Como $\mid Z \mid \geq z_{\alpha/2} = 1.96$ se rechaza la hipótesis nula. Concluimos que ambas poblaciones tienen medias diferentes.

Referencias i

David G. Luenberger.
Linear and Nonlinear Programming.
Addison-Wesley, 1984.

Arthur D. Hall.
Ingeniería de sietemas.
C.E.C.S.A., 1983.

Andrés Ramos, Pedro Sánchez, José María Ferrer, Julián Barquén, and Pedro Linares.
Modelos matemáticos de optimización.

Wayne L. Winston.
Investigación de operaciones. Aplicaciones y algoritmos.
Thomson, 4 edition, 2005.

Referencias ii

R. L. Ackoff.
Planificación de la empresa del futuro.
LIMUSA, 1983.

Juan José Salazar-González.
Programación matemática.
Ediciones Díaz de Santos, 425, 2001.

Enrique Castillo, Antonio J. Conejo, Pablo Pedregal, Ricardo García, and Natalia Alguacil.
Formulación y resolución de modelos de programación matemática en ingeniería y ciencia, 2002.

Abraham Duarte-Mu noz, Juan José Pantrigo-Fernández, and Micael Gallego-Carrillo.
Metaheurísticas.
Librería-Editorial Dykinson, 2007.

S. Koziel and Z. Michalewicz.
Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization.
Evolutionary Computation, 7:pp, 1999.

Zibigniew Michalewicz and David B. Fogel.
How to solve it: modern heuristics.
Springer, 1998.

Local Search in Combinatorial Optimization (Wiley Series in Discrete Mathematics & Optimization).
Wiley, 1997.