

Módulo 3: Midiendo la Exactitud

Carlos López Vázquez

carlos.lopez@ieee.org

Plan

- Error, Exactitud y Precisión
- Definiendo estándares
- Informando sobre la Calidad
- Midiendo y Verificando
- Análisis de Sensibilidad



Error, Exactitud y Precisión

- *Si se lo mide, no se lo puede administrar*
- Tiene implicancias:
 - Sobre el dato final
 - Sobre el proceso productivo
 - Sobre las habilidades de las personas
- Cada dato es específico, pero hay reglas generales

Definiendo estándares

Fuente: Foote and Huebner

- Especificar criterios desde el principio
- Para datos espaciales y no espaciales
- Tópicos a considerar:
 - Niveles de error admisible
 - Diccionarios, tesauros
 - Criterios de clasificación
- Criterios para Datos
- Criterios para Procedimientos
- Tres pasos:

1: qué exigir

- Especificar criterios
 - Mucha exactitud *costa*
 - Poca exactitud *costa*
 - Requisitos son propios del proyecto
- Ir hacia atrás: desde el output a los inputs
- Requerirán Análisis de Sensibilidad

2: entrenar a la gente

- No alcanza con exigir el éxito
- Deben conocerse los objetivos
- Debe capacitarse para lograrlo
- Deben saber *para qué* se hace así

El proyecto no puede salir *a pesar* de la gente

3: verificar procesos y resultados

- Aplique un control regular
- El Control puede ser
 - por lotes
 - continuo
- Aplicarlos
 - A proveedores externos
 - Internamente



Plan

- ✓ Error, Exactitud y Precisión
- ✓ Definiendo estándares
- Informando sobre la Calidad
- Midiendo y Verificando
- Análisis de Sensibilidad



Informando sobre la Calidad

- Siempre debe estar escrito
 - En documentos independientes
 - Dentro de los datos
- Incluir Cómo, Cuándo, Dónde
- Asignar a un responsable
- Documentar preserva la inversión
 - Los datos sobreviven más de lo esperado
 - Típico es 50-100 años
- Y después vendrán los arqueólogos...

Plan

- ✓ Error, Exactitud y Precisión
- ✓ Definiendo estándares
- ✓ Informando sobre la Calidad
- Midiendo y Verificando
- Análisis de Sensibilidad



Midiendo y Verificando

- La realidad manda...
- Comparar en campo si es posible
- Comparar contra otra fuente más precisa si existe
- Existen ensayos específicos por tipo de dato
- Se aplican tests paramétricos y no paramétricos

Cualquiera sea el test...

- Deberá ser científicamente creíble
 - Basado en estadística, medidas o ambas
 - Repetible por otros
- Deberá costar algo razonable
 - Costo en dinero
 - Costo en tiempo
- Ampliamente aceptable
 - Basado en estándares profesionales, nacionales y/o internacionales

Veremos uno...

NSSDA: National Standard for Spatial Data Accuracy (1998)

NSSDA

POSITIONAL ACCURACY HANDBOOK 1999

- Identifica un estadístico bien definido para describir resultados de ensayos de exactitud posicional
- Describe un método para el ensayo de coordenadas o posiciones
- Provee un lenguaje estandarizado para informar la exactitud
- No prevee ensayos para variables temáticas o categóricas; sólo cuantitativas

Siete pasos

1. Determinar si se requiere exactitud horizontal, vertical o ambas
2. Seleccionar un conjunto de test points
3. Identificar y extraer los homólogos de un conjunto independiente de mejor exactitud
4. Tomar medidas en puntos idénticos a los anteriores

Siete pasos (cont)

5. Calcular un estadístico de la exactitud de la posición según corresponda
6. Preparar la declaración de exactitud según el formulario
7. Incluya esa declaración en los metadatos



Siete pasos

1. Determinar si se requiere exactitud horizontal, vertical o ambas
2. Seleccionar un conjunto de test points
3. Identificar y extraer los homólogos de un conjunto independiente de mejor exactitud
4. Tomar medidas en puntos idénticos a los anteriores

Siete pasos

1. Determinar si se requiere exactitud horizontal, vertical o ambas
2. Seleccionar un conjunto de test points
3. Identificar y extraer los homólogos de un conjunto independiente de mejor exactitud
4. Tomar medidas en puntos idénticos a los anteriores

POSITIONAL ACCURACY HANDBOOK 1999

Algunos detalles...

- ¿Dónde localizar los puntos?
- Bien definidos, fáciles de medir
- Intersecciones perpendiculares
- Monumentos
- N>20, fallas ≤1
- ¿N≤20?
 - Estimación deductiva
 - Evidencia interna
 - Comparación con la fuente

Siete pasos

1. Determinar si se requiere exactitud horizontal, vertical o ambas
2. Seleccionar un conjunto de test points
3. Identificar y extraer los homólogos de un conjunto independiente de mejor exactitud
4. Tomar medidas en puntos idénticos a los anteriores

POSITIONAL ACCURACY HANDBOOK 1999

Datos ¿independientes?

- Buscar puntos comunes
- Exactitud
 - Idealmente >> triple
 - En otro caso... lo que haya (¡si es mejor!)
- Documentarla en metadatos
- ¡Usar sólo cifras significativas!
- Exactitud ~uniforme en área
- Área ~equivalente

Siete pasos (cont)

5. Calcular un estadístico de la exactitud de la posición según corresponda
6. Preparar la declaración de exactitud según el formulario
7. Incluya esa declaración en los metadatos

POSITIONAL ACCURACY HANDBOOK 1999

Cálculos en horizontal...

$$RMSE_x = \sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{check,i})^2}{n}}$$

$$RMSE_y = \sqrt{\frac{\sum_{i=1}^n (y_{obs,i} - y_{check,i})^2}{n}}$$

Si: $\frac{RMSE_x}{RMSE_y} \in \left[\frac{6}{10}, \frac{10}{6} \right]$

- Error_x independiente de Error_y
- Errores con distribución Normal
- No hay outliers

→ Exactitud_{NSSDA,x} = 2.4477 * 0.5 * (RMSE_x + RMSE_y)

POSITIONAL ACCURACY HANDBOOK 1999

Cálculos en la vertical...

$$RMSE_z = \sqrt{\frac{\sum_{i=1}^n (z_{obs,i} - z_{check,i})^2}{n}}$$

Si: $\begin{cases} \text{Errores con distribución Normal} \\ \text{No hay outliers} \end{cases}$

→ Exactitud_{NSSDA,z} = 1.9600 * RMSE_z

Variables temáticas o categóricas

- No contempladas en el NSSDA

Es importante hacer foco en:

- Naturaleza de los errores
 - ¿Se confundieron especies de pinos?
 - ¿Se confundió trigo con vid?
- Frecuencia de los errores
- Magnitud de los errores
- Origen de los errores

Algunos serán más graves que otros

© Arthur J. Lembo, Jr. Cornell University

¿Cómo medir la Exactitud?

- Prepare una matriz de confusión
 - Elija N puntos al azar
 - Vea lo que dicen los datos
 - Vea lo que hay en el terreno
 - Cuente los casos
- Idealmente, todo en la diagonal

		Terreno		
		A	B	C
Datos	A	10	2	3
	B	0	20	0
	C	4	1	10

Exactitud global = traza(A)/N

Alguna terminología

- Error por Omisión
 - Def.: Error por columna/total en columna (1.0 - Exactitud *del productor*)
- Error por Comisión
 - Def.: Error por fila/total en fila (1.0 - Exactitud *del usuario*)

¡No incluir los elementos de la diagonal!

Resultados del ejemplo

- Exactitud *Total*:
 $(10+20+10)/(10+2+3+0+20+0+4+1+10) = 40/50 = 80\%$
- Exactitud de *comisión* para la clase A:
 $10/(10+2+3) = 10/15 = 67\%$
- Exactitud de *omisión* para la clase A:
 $10/(10+0+4) = 10/14 = 71\%$

Datos	Terreno		
	A	B	C
A	10	2	3
B	0	20	0
C	4	1	10

La tabla completa

	A	B	C	TOTAL	Exactitud del Usuario
A	10	2	3	15	0.67
B	0	20	0	20	1.00
C	4	1	10	15	0.67
TOTAL	14	23	13	40	
Exactitud del Productor	0.71	0.87	0.77		
Exactitud total	0.8				

El índice *kappa* de Cohen

- Una medida de la exactitud observada en comparación con el azar
- Es un único número (un escalor)
- El ideal sería 1.0; lo peor sería 0.0
- Un valor intermedio 0.83 implica que se están evitando 83% de los errores que cometería una clasificación al azar
- No está recogido en los estándares
- Sin embargo se usa...
- ¿Cómo calcularlo?

Ejemplo:

	A	B	C	TOTAL	Exactitud del Usuario
A	10	2	3	15	0.67
B	0	20	0	20	1.00
C	4	1	10	15	0.67
TOTAL	14	23	13	40	
Exactitud del Productor	0.71	0.87	0.77		
Exactitud total	0.8				

$$q_A = 15 \cdot 14 / 50 = 4.20$$

$$q_B = 20 \cdot 23 / 50 = 9.20$$

$$q_C = 15 \cdot 13 / 50 = 3.90$$

$$q = \text{Total} = 17.3$$

1. Calcular "q" para cada caso
 Def.: q es el número de casos en la diagonal por puro azar
 $q_i = \text{sum}_i(\text{fila}) \cdot \text{sum}_i(\text{col}) / N$

Ejemplo (cont.):

	A	B	C	TOTAL	Exactitud del Usuario
A	10	2	3	15	0.67
B	0	20	0	20	1.00
C	4	1	10	15	0.67
TOTAL	14	23	13	40	
Exactitud del Productor	0.71	0.87	0.77		
Exactitud total	0.8				

$$\text{Kappa} = (40 - 17.3) / (50 - 40) = 0.227$$

2. Calcular *kappa*

$$\text{Def.: } \text{kappa} = (\text{traza}(A) - q) / (N - q)$$

Asistente malo!

Un pequeño problema...

- Se asume que no hay confusión posible al construir la matriz de confusión
- En la práctica...
 - Absolutamente incorrecto
 - Comprensible, pero incorrecto
 - Razonable, pero podría ser mejor
 - Buena respuesta
 - Absolutamente correcto

Digresión: Midiendo Calidad...

- Estamos revisando los siete pasos
- Hemos visto algunos tests de Exactitud
- Algunos estandarizados, otros no
 - Para posición
 - Para atributos
- Esencialmente controlan Exactitud
- Pero en Calidad no todo es Exactitud...

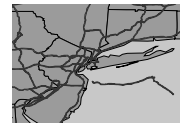
Otros componentes...

- Complejidad
- Coherencia lógica
- Exactitud temporal
- Linaje
- Metadatos

Componentes más modernas
 Sin definir totalmente
 Falta unificar nomenclatura
 Faltan métricas

Ejemplo: Coherencia lógica

- Representación sin sentido
 - Caminos en el agua
 - Isolneas que se cortan
 - Puentes sin rutas o sin río
 - Cuencas hidrográficas vs. ríos



¿Cómo lo mediría?

Volviendo a los Siete pasos...

5. Calcular un estadístico de la exactitud de la posición según corresponda
6. Preparar la declaración de exactitud según el formulario
7. Incluya esa declaración en los metadatos



POSITIONAL ACCURACY HANDBOOK 1999

Declaración de Exactitud

- En el estándar se consideran dos formatos:
 1. “__ mts de exactitud (horizontal/vertical) al 95% de confianza”
 2. “Compilado para tener __ mts de exactitud (horizontal/vertical) al 95% de confianza”
- Usar (1) si se dispone de datos específicos independientes
- Usar (2) si se utilizó un procedimiento estándar consistente con el error obtenido

Incluyendo la Incertidumbre

- Falacia de *Falsa Precisión*
 - Una represa de 1:234.567,89 m³
 - La probabilidad de falla es 0.3579
- Regla (indefendible...): un décimo de la resolución
 - Población en unidades → densidad en décimos
 - Simple y popular, ¡pero errónea!

Incluyendo Incertidumbre⁽²⁾

- Falacia de *Falsa Certidumbre*
 - Respuestas “precisas” sin sustento
 - Exacerbado al combinar capas
 - Ni hablar de mezclar escalas
 - Problema extremadamente corriente
- Lo correcto sería alguno de:
 - resultado-intervalo
 - [peor, típico, mejor]
 - [min,max] con probabilidad
- ¿Cómo determinarlo?

¡Permanezca en este canal!

Siete pasos (cont)

5. Calcular un estadístico de la exactitud de la posición según corresponda
6. Preparar la declaración de exactitud según el formulario
7. Incluya esa declaración en los metadatos



POSITIONAL ACCURACY HANDBOOK 1999

Comunicando...

- Si hay varias capas, podría requerirse una declaración para cada una de ella
- Si hay una mezcla inseparable y no ensayada, asignarle el peor nivel de error
- Si está ensayada, asignarle el nivel obtenido
- Nunca use más cifras que las de los datos

Otras formas de estimar...

- Si tiene pocos puntos
- Si la distribución es claramente no-gaussiana
- ¡No está previsto en el estándar!
 - ➔ Use Remuestreo (*Bootstrap*)

Atención: Lo que de allí resulte complementa, pero no sustituye a lo requerido por el estándar

Breve introducción al Bootstrap

Fuente: Lucila Ohno-Machado

¿Porqué Remuestrear?

- En ocasiones no es posible tener muchas muestras de una población
- En ocasiones no es correcto (o posible) asumir una distribución para una población
- El objetivo es: Assess sampling variation

Bootstrap

- Efron (biostatístico de Stanford) a finales de los 80's
 - "Pulling oneself up by one's bootstraps"
- Es un enfoque no paramétrico para inferencia estadística
- Usa *cálculos (fuerza bruta)* en lugar de resultados asintóticos e hipótesis tradicionales sobre distribuciones
- Puede usarse para obtener errores estándar, intervalos de confianza y prueba de hipótesis

Ejemplo

- Adaptado de Fox (1997) "Applied Regression Analysis"
- Objetivo: Estimar diferencia promedio de respuestas entre Hombre y Mujer
- Se dispone únicamente de cuatro parejas de observaciones:

Observ.	Hombre	Mujer	Differ.
1	24	18	6
2	14	17	-3
3	40	35	5
4	44	41	3

Diferencia promedio

- Promedio \bar{Y} de la muestra es $(6-3+5+3)/4 = 2.75$
- Si Y fuese normal, el Intervalo de Confianza CI al 95% sería

$$\mu = \bar{Y} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$
- *Problema:* no conocemos σ

Estimadores corrientes

- El estimador de σ es $S = \sqrt{\frac{\sum (y_i - \bar{Y})^2}{(n-1)}}$
- El estimador de la desviación estándar es $SE(\bar{Y}) = \frac{S}{\sqrt{n}}$
- Asumiendo que la población es normal, podemos usar la distribución-t como $\mu = \bar{Y} \pm t_{n-1, 0.025} \frac{S}{\sqrt{n}}$

Intervalo de Confianza CI

$$\mu = \bar{Y} \pm t_{n-1, 0.025} \frac{S}{\sqrt{n}}$$

$$\mu = 2.75 \pm 4.30 (2.015) = 2.75 \pm 8.66$$

$$-5.91 < \mu < 11.41$$

WTF!!!

Media y varianza de la muestra

Idea: usar la distribución Y^* de la muestra para estimar la distribución Y de la población

y^*	$p^*(y^*)$	
6	.25	
-3	.25	$E^*(Y^*) = \sum y^* p(y^*) = 2.75$
5	.25	$V^*(Y^*) = \sum [y^* - E^*(Y^*)]^2 p(y^*) = 12.187$
3	.25	$DE^*(Y^*) = \sqrt{V^*(Y^*)} = 1.745$

Muestreo con reposición

Caso extremo

Muestra	Y_1^*	Y_2^*	Y_3^*	Y_4^*	\bar{Y}^*
1	6	6	6	6	6.00
2	6	6	-3	-3	3.75
3	6	6	5	5	5.75
...					
100	-3	5	6	3	2.75
101	3	3	3	3	3.00
...					
255	-3	3	3	5	3.5
256	3	3	3	3	3.00

Sólo hay 256 posibilidades

Otro caso extremo

Calculando nuevos CI

El promedio de estos 256 promedios es nuevamente 2.75, pero su desviación es ahora

$$DE^*(\bar{Y}^*) = \sqrt{\frac{\sum (\bar{Y}_k^* - \bar{Y}^*)^2}{n}} = 1.745$$

(sin "A" porque la desviación ahora no se estima)

$$DE(\bar{Y}) = \sqrt{\frac{n}{n-1}} DE^*(\bar{Y}^*)$$

$$2.015 = \sqrt{\frac{4}{3}} \times 1.745$$

WTF!!!

¿Y entonces?

- ¡Esto ya era sabido!
- Pero mediante remuestreo o bootstrap
 - Los Intervalos de Confianza pueden ser más precisos
 - Pueden ser calculados para problemas no lineales que no tienen fórmulas de error conocidas

**La población es a la muestra
lo mismo que
la muestra es a las remuestras**

En la práctica (a diferencia del ejemplo anterior), *no todas* las N^N remuestras son seleccionadas

Procedimiento

1. Especifique un criterio que produzca la muestra que utilizará
 - Leáse como "el criterio aplicado a la población"
 - Criterio(población) → muestra
2. Use esta muestra como si fuese la población (pero *con reemplazo*)
 - Criterio(muestra) → re-muestra1
re-muestra2
etc...

Procedimiento⁽²⁾

3. Para cada remuestra, calcule el estimador estadístico de su interés
4. Use la distribución de los estimadores de las remuestras para estimar las propiedades de la muestra

Atención:

- Remuestreo no es válido *absolutamente* para todo
- No debe usarse si hay colas "largas"
- Ej.: estimación del rango

<http://www.itl.nist.gov/div898/handbook/eda/section3/bootplot.htm>

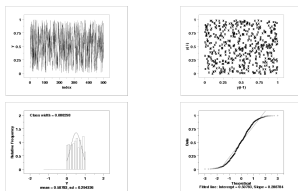
Otro ejemplo:

- Ver si unos datos ajustan al modelo

$$Y_i = C + e_i$$
- Confirmar que el proceso está *bajo control*
 - Tomar muestras al azar y confirmar que es de distribución, parámetros y escala fijos
- En particular, confirmar CI al 95%

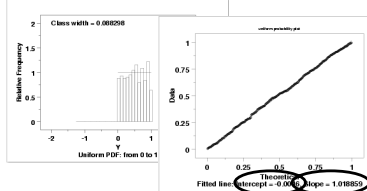
$$CI = \left[\bar{Y} - 2.0 \frac{s}{\sqrt{n}}, \bar{Y} + 2.0 \frac{s}{\sqrt{n}} \right]$$

¿Qué hacer primero? 4 gráficos



No luce como normal...

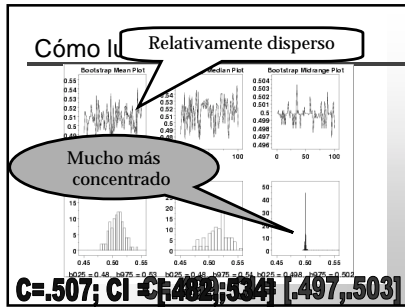
Supongamos Uniforme...



Ahora sí...


Un estimador para C

- Y_i no es normal
- Aplicaremos remuestreo
- Estimadores posibles de C
 - Media
 - Mediana
 - Promedio del rango
 - Quizá otros...
- Buscaremos el de mínima varianza



Importancia...

- Lo típico es requerir CI para la media
 - Caso normal: hay solución analítica
- En otros casos no la hay
 - Es donde el bootstrap/resampling muestra su importancia



Otros métodos de remuestreo


- Jackknife (*sacar de a uno*) es un caso especial de bootstrap
 - Remuestreo sin un caso y sin reemplazo (las muestras pasan a tener tamaño $n-1$)
- Validación cruzada
 - Divide los datos en entrenamiento y test
- Generalmente se los usa para estimar intervalos de confianza en predicciones para el modelo "full" (i.e., modelo que usa todos los casos)

Recapitulando:

- Vimos estándares
- Vimos métricas de Exactitud
 - Posicionales
 - Atributos Cuantitativos, Cualitativos, etc
- Vimos métodos para medir Exactitud
 - Cuando hay acceso a *suficientes* valores *mejores*
 - Cuando son gaussianos
- Para los otros casos
 - Explicamos e ilustramos Bootstrap
 - Mencionamos Jackknife

Plan

- ✓ Error, Exactitud y Precisión
- ✓ Definiendo estándares
- ✓ Informando sobre la Calidad
- ✓ Midiendo y Verificando
- Análisis de Sensibilidad



Análisis de Sensibilidad

- Siempre hay que Validar los resultados
 - Conjunto de entrenamiento
 - Conjunto de validación
- ¿Cómo? Depende del Modelo+Datos
- Datos continuos → derivada parcial
- Datos categóricos → simulación
- Se suele hacer al final... ¡pero puede invalidar todo el proyecto!

Cuanto antes mejor...

- Si el dato no existe con la exactitud requerida
- Si no era necesaria tanta exactitud ¡Todo cuesta!
- Datos pueden ser irrelevantes...
 - Tipo de suelo para estudio eólico
- Relevancia diferente según el output
 - Producción de energía en kWh-año
 - Costo del proyecto en USS

Ejemplo analítico 1

- Supongamos $f(x) = ax + b$
- ¿Cómo se relaciona la Var(f) con Var(x)?

$$V(f) = \langle f^2 \rangle - \langle f \rangle^2$$

$$= \langle (ax+b)^2 \rangle - \langle ax+b \rangle^2$$

$$= a^2 \langle x^2 \rangle + 2ab \langle x \rangle + b^2 - a^2 \langle x \rangle^2 - 2ab \langle x \rangle - b^2$$

$$= a^2 (\langle x^2 \rangle - \langle x \rangle^2)$$

$$= a^2 V(x) \leftarrow \text{i.e. } \sigma_f = |a| \sigma_x$$
- Más en general:

$$V(f) = \left(\frac{df}{dx} \right)^2 V(x) ; \sigma_f = \left| \frac{df}{dx} \right| \sigma_x$$

Pero vale sólo si la *aproximación lineal* es buena en el rango del error

Ejemplo analítico 2

- Consideremos ahora $f = ax + by + c$

$$V(f) = a^2 \langle x^2 \rangle - \langle x \rangle^2 + b^2 \langle y^2 \rangle - \langle y \rangle^2 + 2ab \langle xy \rangle - \langle x \rangle \langle y \rangle$$

$$= a^2 V(x) + b^2 V(y) + 2ab \text{cov}(x, y)$$
- En general

$$V(f) = \left(\frac{df}{dx} \right)^2 V(x) + \left(\frac{df}{dy} \right)^2 V(y) + 2 \left(\frac{df}{dx} \right) \left(\frac{df}{dy} \right) \text{cov}(x, y)$$

$$\sigma_f^2 = \left(\frac{df}{dx} \right)^2 \sigma_x^2 + \left(\frac{df}{dy} \right)^2 \sigma_y^2 + 2 \left(\frac{df}{dx} \right) \left(\frac{df}{dy} \right) \rho \sigma_x \sigma_y$$

Nuevamente sólo es válida si la aproximación lineal es buena en el rango del error

El coeficiente de correlación $\rho \in [-1, 1]$ es 0 si x,y no están correlacionadas

Otras fórmulas analíticas

- Ahora considérese $f = x \cdot y$ (solo creámoslo...)
 $V(f) = y^2 V(x) + x^2 V(y)$
- $\left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$
- Resultados similares para $f = x/y$
- Otras fórmulas útiles
 $\frac{\sigma_{1/x}}{1/x} = \frac{\sigma_x}{x}$; $\sigma_{\ln(x)} = \frac{\sigma_x}{x}$
El error relativo en x o 1/x es el mismo. Error en el logaritmo es el error relativo.

En general...

- Hay que estimar las derivadas parciales
- Métodos numéricos *aproximados*:
 - Cociente incremental
 $\frac{\partial f}{\partial x_i} \approx \frac{f(x_1, x_2, \dots, x_i + \Delta, \dots, x_n) - f(x_1, x_2, \dots, x_i, \dots, x_n)}{\Delta}$
 - Uso de números complejos
 $\frac{\partial f}{\partial x_i} = \lim_{\Delta \rightarrow 0} \frac{\text{Im}(f(x_1, x_2, \dots, x_i + \Delta^j, \dots, x_n))}{\Delta}, j = -1$
- Problema: hay que estimar incremento Δ apropiado
- Alternativa: uso de la derivada exacta

Derivada exacta...

- Derivación manual
 - Sólo en casos simples; requiere codificar
- Derivación automática
 - Variante 1: Generador de nuevo código
 - Variante 2: Sobrecarga de operadores
- Paquetes
 - ADOL-F/ADOL-C
 - TAPENADE
 - Matlab+ADMAT
 - Otros...
- Ejemplos



Generadores de código

- Generan automáticamente un segundo código fuente a partir del disponible
- ADOL-C/ADOL-F/TAPENADE, etc.
- Derivada exacta; no hay cociente incremental
- Ejemplo:
 - $v_2 = 2 * v_1 + 5$
 - $\dot{v}_2 = 2 * \dot{v}_1$
 - $v_4 = v_2 + p_1 * v_3 / v_2$
 - $\dot{v}_4 = \dot{v}_2 * (1 - p_1 * v_3 / v_2^2) + \dot{v}_3 * p_1 / v_2$

Sobrecarga de operadores

- Sólo en ambientes orientados a objetos
 - Ej: Matlab + ADMAT toolbox
- ```
function y = getfun(x)
 z = x*x;
 w = x + z;
 y = w*w
```
- ```

x = 2.0
dx/dt = 1.0
-----
z = 4.0
dx/dt = 2*x*dx/dt = 4.0
-----
w = 6.0
dx/dt = dx/dt + dx/dt = 1.0
-----
y = 36.0
dx/dt = 2*w*dx/dt = 6.0
    
```
- Ventajas:
 - Derivada exacta (tampoco hay cociente incremental)
 - Sólo hay un código fuente

¿En qué andábamos?

- Estábamos haciendo un Análisis de Sensibilidad
 - ¿Alcanza/no alcanza la Exactitud de los datos disponibles?
- Vimos Métodos Analíticos
 - Sólo para datos continuos
- Veremos métodos Estadísticos de Simulación
 - Para datos continuos/catórgicos



Métodos de Monte Carlo

Fuente: Vesna Lizar-Stiffler

- Cualquier procedimiento que usa números aleatorios
- ¿Números aleatorios?
 - ¿Qué es un número aleatorio?
 - ¿Un número?
 - Secuencia de números aleatorios



La parte aleatoria...

- Distribución Uniforme
 - Todo número en [0,1] tiene la misma probabilidad
 - Fundamental (se usa para otras distribuciones)
 - Para muestreo al azar, etc.
- ¿Cómo generar una secuencia de números aleatorios?
 - Usar un sistema caótico (moneda, dado, etc.)
 - Usar un proceso inherentemente randómico
 - Usar tablas de números aleatorios
- Problema:
 - Resultados no reproducibles
 - Impráctico, no confiable (¿sesgos?)
 - Tablas - no suficientemente extensas

Números pseudo-aleatorios

- Números pseudo-aleatorios generados con algoritmos
 - No correlacionados, ciclos largos, etc.
- Ej: Middle square algorithm (J. Von Neumann, 1946):
 - Para generar una secuencia de enteros de 10 dígitos:
 - Elige uno cualquiera.
 - Elevalo al cuadrado y luego
 - Extraiga los 10 dígitos intermedios como el siguiente número de la secuencia
 - Por ejemplo:

$$3690295441^2 \quad 2$$

$$13618280441865384481$$

Seudo-Aleatorios...

- ¡No son al azar! Son series totalmente determinísticas y predecibles
- Lucen como *al azar*, pero ...

$$\begin{array}{r} 6100^2 \\ \hline 3721000 \\ \hline 441000 \\ \hline 1461000 \\ \hline 6561000 \end{array}$$

¡LOOP!

Linear congruential method (Lehmer, 1948)

- $I_{n+1} = (a \cdot I_n + c) \bmod m$
 - $I_0 =$ Valor inicial (semilla)
 - $a, c \geq 0$,
 - $m > I_0, a, c$
 - Una mala elección de constantes generará secuencias malas:
 - Ej.: $a=c=I_0=7, m=10 \rightarrow$
 - 7, 6, 9, 0, 7, 6, 9, 0, ...
- A mod B=resto de dividir A entre B

Familia popular...

- RANDU: $I_{n+1} = (65539 \cdot I_n) \bmod 2^{31}$
 - Se vio que no era bueno
- RANMAR: $I_n = (a \cdot I_{n-1} + b \cdot I_{n-2}) \bmod m$
 - genera series con periodo $\approx 10^{43}$
- En general se deberían usar generadores conocidos, con propiedades bien documentadas
- Sin embargo...

Para PDF's no uniformes

- Puede haber rutinas específicas (ej.: $N(0,1)$)
- ¿Otros casos?
 - Método de la probabilidad inversa
 - Calcular $F(x) = \int_{-\infty}^x f(s) ds$
 - Generar $t \sim U(0,1)$
 - Evaluar $x = F^{-1}(t)$
 - x resulta tener pdf igual a $f(x)$
 - Hay otros métodos para casos particulares
- Problemas: pdf's discretas, eficiencia, etc.

Simulación de Monte Carlo

- Aparece con las primeras computadoras (1945-55)
- Áreas de aplicación incluyen biología, química, informática, análisis de datos econométricos y financieros, ingeniería, ciencia de los materiales, física, ciencias sociales, estadística, etc.
- La SMC se resume en:
 - Asumiendo un mecanismo para generar datos de su proceso
 - Produzca nuevas instancias de datos simulados
 - Examine estadísticamente los resultados de esas instancias
- Si se tiene "la" pdf \rightarrow SMC estándar
- Si no, se usa Remuestreo (que es un pariente cercano)

Típicamente medias y varianzas

Algunos problemas...

- SMC plantea gran demanda de CPU
 - Varianza $\sim 1/\sqrt{N}$
 - Bajar el error un 10% requiere multiplicar por 100 el número de simulaciones
- Se intenta mejorar esta característica
 - Técnicas de Reducción de Varianza (TRV)
- Veamos de qué se trata...

TRV: Enfoque Antitético

- Se basa en que

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2)$$
- Si X_1 y X_2 son independientes,

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2)$$
- Si se pueden generar X_1 y X_2 de forma que tengan correlación $\text{Cov}(X_1, X_2)$ negativa puede lograrse que $\text{Var}(X_1 + X_2)$ sea *menor* que si son independientes

Ejemplo: Integral (o promedio)

$$\hat{I} = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \text{donde } X_1, X_2, \dots, X_{2n} \stackrel{iid}{\sim} f(x)$$

Puede ser reformulado como

$$\hat{I} = \frac{1}{n} \sum_{i=1}^n Y_i \quad \text{donde } Y_i = \frac{X_i + X'_i}{2}$$

donde X_i y X'_i están correlacionadas *antitéticamente* y las Y_i siguen siendo independientes

¿Cómo lograr correlación negativa?

- El procedimiento en general es

$$X = F^{-1}(U) \quad \text{donde } U \sim U(0,1) \text{ y } dF/ds = f(s)$$
- Pero $U \sim U(0,1)$ implica que $(1-U) \sim U(0,1)$.
- Por lo que, $X' = F^{-1}(1-U)$ también tiene la misma distribución $f(s)$ que X
- Si $F(s)$ es monótona (y por lo tanto también $F^{-1}(s)$) entonces X' y X tienen correlación negativa

Otros trucos...

■ Si se quiere calcular $I = \int_a^b g(x) dx$

■ En ocasiones es fácil resolver analíticamente la integral

$$I' = \int_a^b h(x) dx$$

■ y posteriormente hacer la simulación de Monte Carlo en el residuo, logrando

$$I = I' + \int_a^b [g(x) - h(x)] dx$$

MonteCarlo

Aspectos importantes en SMC

- Diseño apropiado
- Medida de la exactitud del resultado
- Justificación del número de casos/muestras
- Selección del generador aleatorio
- Memoria y tiempo de cálculo
- Técnica de reducción de varianza
- Software usado
- Análisis del resultado (visualización, ajuste por curvas, etc.)



En resumen:

- Como productor: hay estándares para expresar Exactitud
 - Para posición y variables continuas
 - No para variables categóricas
- Estimación basada en suficientes valores más exactos
 - Alternativamente, usar Bootstrap
- Como usuario: preocuparse *mucho* sólo si el problema es *sensible*

¿Cómo medir Sensibilidad?

■ Hay técnicas informáticas...

➢ Simulación Monte Carlo

- Uso general, no paramétrica, etc.
- Mucha CPU y tiempo

➢ Propagación de errores

- Sólo funciones diferenciables
- No requiere simulación
- Complejas para algunas operaciones GIS típicas

■ Escasa disponibilidad en GIS

Módulo 3: Midiendo la Exactitud

Carlos López Vázquez

carlos.lopez@ieee.org