

# Simulación IV

Dpto. Ingeniería Industrial, Universidad de Chile

IN47B, Ingeniería de Operaciones

# Contenidos

- 1 Reduccion de Varianza
- 2 Analizando Resultados
- 3 Otros aspectos de Simulación

# Método

Para estimar  $\mu$  tenemos un simulador que entrega variables aleatorias

- $X$  tal que  $\mathbb{E}(X) = \mu$
- $Z$  con  $\mathbb{E}(Z)$  conocido

$$\forall c \quad X_c = X + c(Z - \mathbb{E}(Z)) \text{ tal que } \mathbb{E}(X_c) = \mu$$

Estimadores insesgados de  $\mu$

- $\bar{X}$
- $\bar{X}_c = \bar{X} + c(\bar{Z} - \mathbb{E}(Z))$  para todo  $c$

# Método

Donde tenemos que

$$\mathbb{V}\text{ar}(X) = \mathbb{V}\text{ar}(X) + c^2 \mathbb{V}\text{ar}(Z) + 2c \text{Cov}(X, Z)$$

El mejor valor de  $c$  es

$$c^* = -\frac{\text{Cov}(X, Z)}{\mathbb{V}\text{ar}(Z)}$$

con lo que se obtiene una minima varianza de

$$\mathbb{V}\text{ar}(X_c) = \mathbb{V}\text{ar}(X) - \frac{\text{Cov}(X, Z)^2}{\mathbb{V}\text{ar}(Z)}$$

# Uso en simulación

Este método

- reduce varianza si  $Cov(X, Z) \neq 0$
- valido en gral. dado
  - estimador insesgado  $X$
  - $Z$  el 'control variate' de  $X$ .

Si un simulador arroja  $X_i$  y  $Z_i$  con  $\mathbb{E}(Z)$  conocido, tenemos el estimador insesgado

$$\frac{1}{n} \sum_{i=1}^n (X_i + c^*(Z_i - \mathbb{E}(Z)))$$

¿que falta?

# Uso en simulación

$$c^* = -\frac{\text{Cov}(X, Z)}{\text{Var}(Z)}$$

- hacer  $p$  corridas piloto y estimar
- $\text{Cov}(X, Z) \sim \frac{1}{p-1} \sum_{j=1}^p (X_j - \bar{X}(p))(Z_j - \bar{Z}(p))$
- $\text{Var}(Z) \sim \frac{1}{p-1} \sum_{j=1}^p (Z_j - \bar{Z}(p))^2$

# Ejemplo

Queremos estimar  $\mathbb{E}(e^{(U+W)^2})$  con  $U, W \sim U[0, 1]$  independientes.

Posibles control variates

- $Z_1 = U + W$
- $Z_2 = (U + W)^2$

Note que  $\mathbb{E}(Z_1) = 1$  y  $E(Z_2) = 7/6$

# Ejemplo

## Matlab Code for Estimating $\theta = E[e^{(U+W)^2}]$

```
> % Do pilot simulation first
> p=100; n=1000;
> u=rand(p,1); w=rand(p,1); y=exp((u+w).^2); z2 = (u+w).^2;
> cov_est=cov([y z2])

cov_est =
    31.3877    4.1800
    4.1800    0.6960

> c_est = - cov_est(1,2)/cov_est(2,2)

c_est = -6.0061

> % Now do main simulation
> u = rand(n,1); w=rand(n,1); y = exp((u+w).^2); z2 = (u+w).^2;
> v = y + c_est*(z2 - 7/6);
> mean(y), std(y)      % Check the mean and std of usual estimator

ans =    5.0340    6.2507

> mean (v), std(v)      % Check the mean and std of new estimator

ans =    4.9329    3.1218

> % Now compute confidence intervals
> CI = [mean(v) - 1.96*std(v)/sqrt(n), mean(v) + 1.96*std(v)/sqrt(n)]

CI =    4.7394    5.1263
```



# El Problema

- En general queremos comparar distintas configuraciones.
- Significa estimar parametros y compararlos.
- ¿Cuándo podemos decir que son distintos?

## Ejemplo

Compararemos un sistema M/M/1 con un sistema M/M/2. En sistema M/M/1 llegadas son 10 por unidad de tiempo, y atendemos 11 clientes por minutos. En sistema M/M/2 llegadas son 10 por unidad de tiempo, y cada servidor atiende 5.5 clientes por minutos.

# Comparando estimadores de $\mu$

- Supongamos que ambas configuraciones cuestan lo mismo.
- Escoger sólo dependera de calidad de servicio.
- Medimos calidad de servicio como tiempo espera promedio.

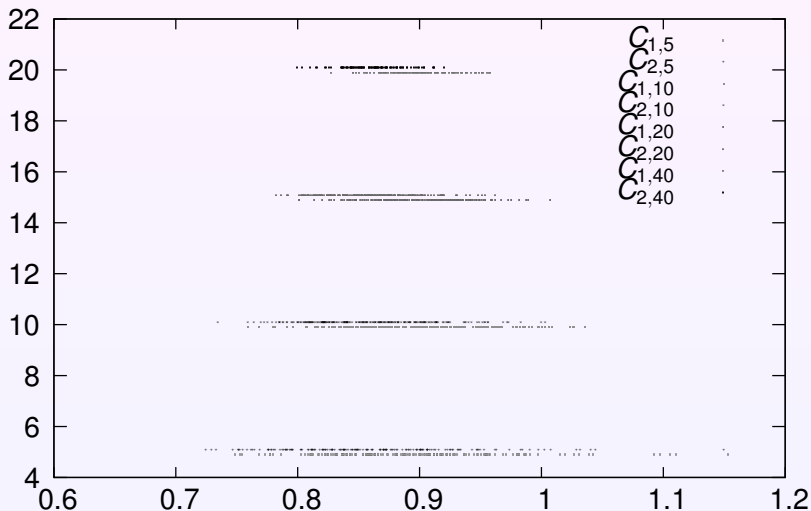
## Primer reflejo

Supongamos que tenemos un *especialista* que sabe de simulación. Simula ambos sistemas, y computa una estimación de  $\mu$  para ambos sistemas con  $k$  corridas independientes.

El decide escoger el sistema con mejor tiempo de espera *estimada*

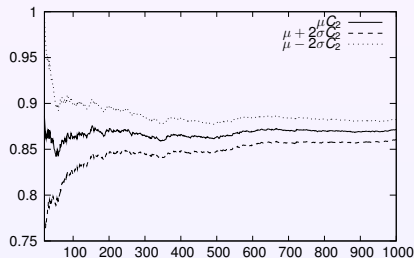
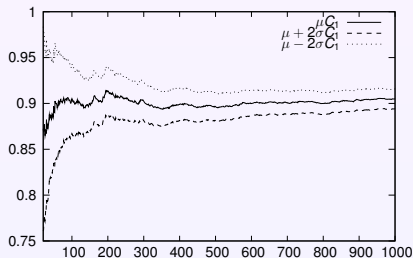
## Comparando Distintas Configuraciones

## ¿Cómo nos va?



# Otro Enfoque

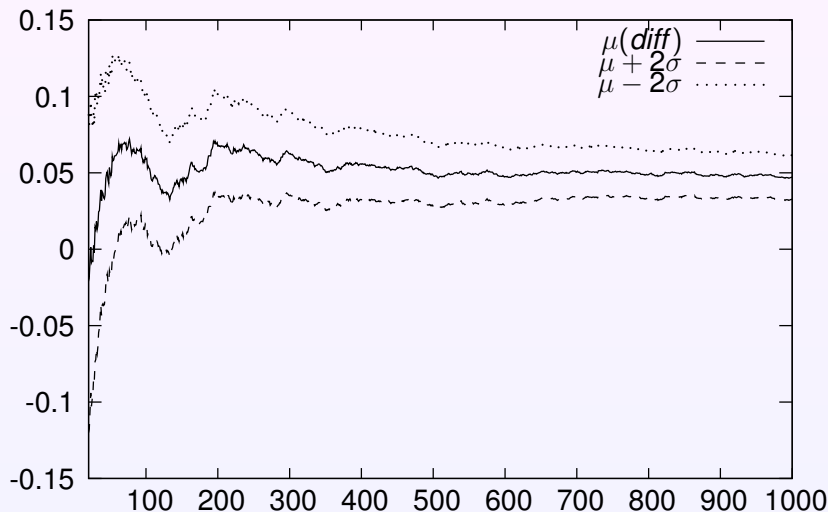
- Un mejor enfoque seria calcular intervalos de confianza para ambos  $\mu$ .
- Habria que alcanzar un nivel donde ambos intervalos no se overlapan.



# Algunas Observaciones

- El enfoque anterior funciona.... pero
- Típicamente requiere un número de replicaciones altas.
- Sufre de los problemas de no-simetria de las distribuciones subyacentes.
- El problema anterior podríamos solucionarlo.
- Considerando diferencia de los estimadores.
- Analizamos  $Z_n = X_n - Y_n$ .
- $Z_n$  tiende a tener una alta simetria.
- Decimos que configuraciones son distintas si 0 no esta en intervalo de confianza.

# ¿Cómo nos va con este enfoque?



# Podemos Mejorar?

- Otra forma de comparar sistemas o configuraciones.
- Exacerbar diferencias en lo sistemas.
- Comparar bajo situaciones de stress del sistema.
- Ello conlleva a diferencias más sustanciales en los estadísticos.

# Caso de múltiples configuraciones:

## Comparación con Standard:

- Elegimos una configuración base  $X^0$ .
- Comparamos  $k$  configuraciones  $X^i : i = 1, \dots, k$ .
- Computamos intervalos de confianza de  $X^0 - X^i : i = 1, \dots, k$  a nivel  $1 - \alpha/k$ .
- Obtenemos Intervalo de confianza global de  $1 - \alpha$ .



# Caso de múltiples configuraciones:

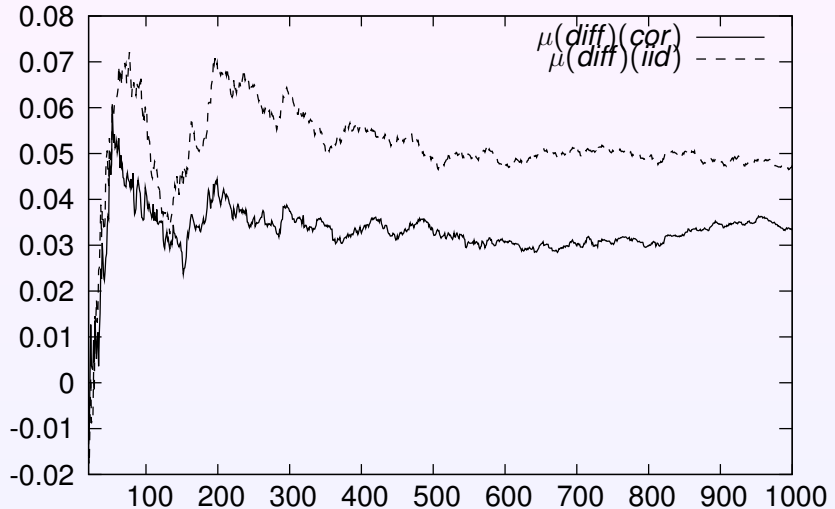
## Comparacion todos los pares:

- Dado  $k$  configuraciones  $X^i : i = 1, \dots, k$ .
- Comparamos todos los pares  $i, j$ .
- Computamos intervalos de confianza para  $X^i - X^j : i, j = 1, \dots, k, i \neq j$  de nivel  $1 - \alpha/k(k-1)$ .
- Obtenemos Intervalo de confianza global de  $1 - \alpha$ .

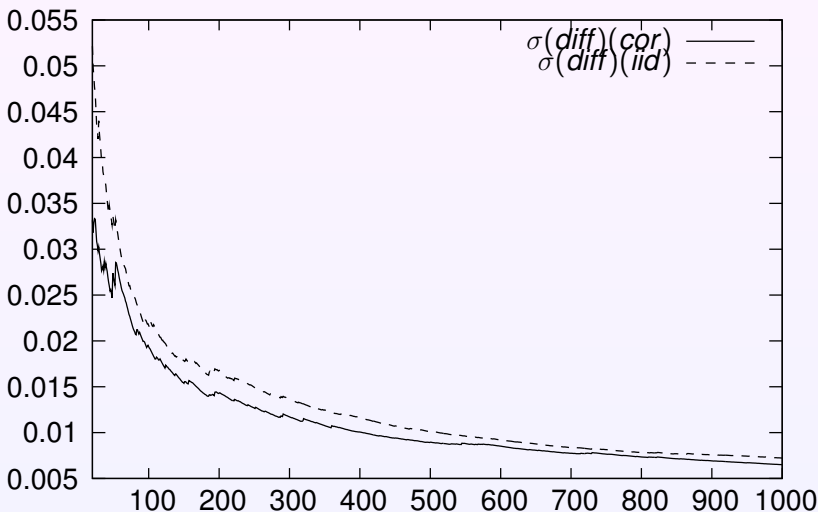
# ¿Qué comparamos cuando comparamos?

- Típicamente comparamos un sistema bajo distintas reglas de operación.
- Los enfoques anteriores comparan ciegamente.
- Llevan a una alta varibilidad.
- ¿Qué queremos realmente?
- Comparar sistema en condiciones lo más cercanas posibles.
- En nuestro ejemplo, ¿qué significaría esto?
- Sistema bajo *misma* demanda.
- Notese que ahora correlación de  $X$  e  $Y$  no es cero.
- Lleva a menores varianzas.
- requiere menos replicaciones para resultados confiables.

# Impacto en Ejemplo



# Impacto en Ejemplo



¿Qué quedó afuera?

# ¿Qué quedó afuera?

- Generar variables aleatorias correlacionadas.
- Validando supuestos del modelo
  - Habría que hacer análisis estadísticos.
- Caso de sistemas en estado estacionario.
- Análisis de sistemas oscilantes.
- Buscando buenos generadores de números aleatorios.