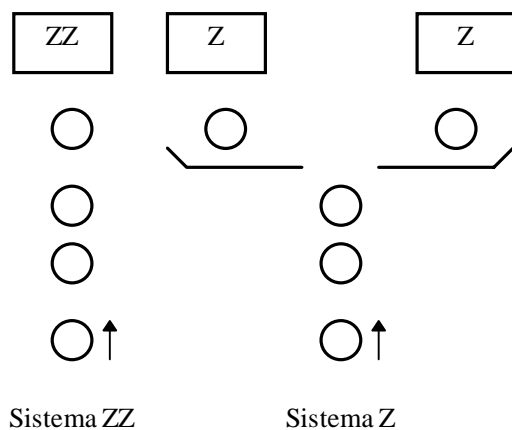


## 8.- Técnicas estadísticas para comparar sistemas alternativos.

En el apartado anterior hemos discutido cómo analizar la información de salida de un único modelo de simulación. Ahora veremos cómo analizar la salida de diferentes modelos de simulación que representan diseños o políticas alternativas.

Ejemplo: un banco desea automatizar cierto servicio al cliente y debe decidir si comprar una máquina modelo ZZ o, por el contrario, dos máquinas modelo Z. La máquina ZZ tiene un precio de compra y mantenimiento doble que la Z, pero es el doble de rápida. Dado que el coste para el banco es independiente de la decisión, desea instalarse el sistema que preste un mejor servicio. Se ha estimado que los clientes llegan de uno en uno, de acuerdo a un proceso de Poisson de frecuencia 1 cliente por minuto. Los tiempos de servicio de los modelos ZZ y Z son variables aleatorias exponenciales IID de media 0.9 y 1.8 minutos respectivamente, independientes del proceso de llegada. Si se instalan las dos máquinas Z, se formará una única cola frente a ambas (ver figura).



La medida de interés es el tiempo medio de espera por cliente en la cola de los 100 primeros clientes, suponiendo que el primer cliente al llegar encuentra el sistema desocupado.

Un método rudimentario de abordar el problema consiste en realizar un único experimento: una única simulación de longitud 100 clientes para cada sistema (empleando secuencias de números aleatorios diferentes para cada uno), y usar el promedio de los 100 tiempos de espera en cada caso para decidir qué instalación conduce a un promedio de espera menor. Sin embargo, el método deja una pregunta sin respuesta: ¿cual es la probabilidad de que el análisis de la recomendación correcta?.

Un método algo menos rudimentario consiste en realizar 100 experimentos completos independientes y anotar cuantas veces la simulación aconseja la instalación ZZ y cuantas la Z. Así, por ejemplo, si obtenemos que en 56 de los 100 experimentos el tiempo medio de espera del sistema ZZ era superior al del sistema Z, puede estimarse, con un 44% de probabilidad de error, que el sistema Z es preferible. Sin embargo, se ha fijado el número de experimentos de forma arbitraria, con lo cual cabe preguntarse si diferiría mucho el resultado obtenido si en vez de haber realizado 100 experimentos se hubieran realizado, por ejemplo, 1000. Como vemos, hay que emplear métodos más sofisticados.

Un requerimiento básico para usar métodos estadísticos en la comparación de sistemas alternativos, es que las observaciones recogidas sean IID y que sus valores esperados sean iguales a las magnitudes que se desean estimar. Para simulaciones con terminación esto se consigue sencillamente realizando replicaciones independientes. Sin embargo, si se desean estimar características estacionarias del sistema la situación no es tan sencilla, ya que no se pueden obtener observaciones IID con valor esperado igual a la respuesta deseada en el estacionario. Un método puede ser emplear la aproximación de las medias por lotes para obtener observaciones IID no sesgadas.

## 8.1- Comparación entre dos sistemas

El *método de los emparejamientos* es útil para comparar dos sistemas, en base al valor esperado o medio de alguna magnitud de interés, mediante la construcción de un intervalo de confianza para la diferencia de los dos valores esperados.

Supongamos que  $X_{i1}, X_{i2}, \dots, X_{in_i}$  (con  $i=1,2$ ) son  $n_i$  observaciones IID del sistema  $i$ , obtenidas en  $n_i$  replicaciones de la simulación. Por ejemplo,  $X_{1j}$  puede ser el promedio de los 100 tiempos de espera en la  $j$ -ésima repetición independiente de la simulación del sistema ZZ.

Sea  $\mu_i = E(X_{ij})$  la magnitud media de interés del sistema  $i$ . Deseamos construir un intervalo de confianza para  $\zeta = \mu_1 - \mu_2$ . Emparejando las observaciones  $X_{1j}, X_{2j}$  (para ello, si  $n_1 \neq n_2$ , deben despreciarse datos del conjunto de observaciones mayor), se definen sus diferencias dos a dos:

$$Z_j = X_{1j} - X_{2j} \quad \text{para } j=1,2,\dots, n = \min\{n_1, n_2\}$$

Las variables aleatorias  $Z_j$  son IID. Un intervalo de confianza para  $E(Z_j) = \zeta$  del  $100(1-\alpha)\%$  es:

$$\bar{Z} \pm t_{n-1, 1-\frac{\alpha}{2}} \sqrt{\hat{\sigma}^2[\bar{Z}]}$$

donde:

$$\bar{Z} = \frac{\sum_{j=1}^n Z_j}{n} \quad \hat{\sigma}^2[\bar{Z}] = \frac{\sum_{j=1}^n [Z_j - \bar{Z}]^2}{n(n-1)}$$

Si los  $Z_j$ 's están distribuidas normalmente, entonces el intervalo de confianza es exacto: contiene a  $\zeta$  con probabilidad  $(1-\alpha)$ . En caso contrario, el teorema del límite central establece que, para  $n$  grande, la probabilidad de que el intervalo contenga a  $\zeta$  es aproximadamente  $(1-\alpha)$ .

Obsérvese que no ha sido preciso imponer que  $X_{1j}$  y  $X_{2j}$  sean independientes, ni que  $Var(X_{1j})$  y  $Var(X_{2j})$  sean iguales.

El permitir una correlación positiva entre  $X_{1j}$  y  $X_{2j}$  es de gran importancia, ya que, como veremos, conduce a una reducción en  $Var(Z_j)$  y, por tanto, a una reducción de la anchura del intervalo de confianza.

**Ejemplo.** Supongamos que, en el modelo del inventario, deseamos comparar dos políticas (s,S) diferentes, (20,40) y (20,80), en términos de su efecto sobre el coste total promedio mensual durante los primeros 120 meses de operación, asumiendo que inicialmente hay inventariadas 60 unidades de producto.

En este caso,  $X_{ij}$  es el coste mensual medio de la política  $i$  ( $i=1,2$ ) en la repetición independiente  $j$ -ésima de la simulación. Realizamos  $n_1 = n_2 = n = 10$  repeticiones independientes de la simulación para cada política, obteniendo los siguientes resultados:

j	1	2	3	4	5	6	7	8	9	10
$X_{1j}$	129.35	127.11	124.03	122.13	120.43	118.38	130.17	129.77	125.51	133.75
$X_{2j}$	127.39	122.52	123.08	119.19	124.46	118.04	123.67	121.59	127.38	119.13

Emparejando las observaciones obtenemos:

$$\bar{Z} = 3.42$$

$$\hat{\sigma}^2[\bar{Z}] = 2.89$$

obteniéndose el siguiente intervalo de confianza del 90% :  $[0.30, 6.53]$  para  $\zeta = \mu_1 - \mu_2$ . Así pues, con un 90% de confianza, podemos afirmar que  $\mu_1 > \mu_2$ , con lo cual la política 2 es preferible a la 1, ya que conduce a un coste promedio menor.

## 8.2- Comparación entre más de dos sistemas

Consideremos ahora el caso en que deba escogerse, de entre  $k \geq 2$  sistemas, aquel que sea el "mejor" de acuerdo a cierta magnitud de interés.

Sea  $X_{ij}$  la variable aleatoria de interés, obtenida de la repetición independiente  $j$ -ésima de la simulación del sistema  $i$ -ésimo.

Sea  $\mu_i = E(X_{ij})$ , con  $i=1, \dots, k$ . Las variables aleatorias  $X_{i1}, X_{i2}, \dots$  se suponen IID, y las simulaciones para diferentes sistemas deben hacerse unas independientes de otras. Por ejemplo,  $X_{ij}$  podría ser el coste medio mensual obtenido en la  $j$ -ésima repetición de la simulación de la política de inventario  $i$ -ésima.

Sea  $\mu_{i_l}$ , la  $l$ -ésima más pequeña de las  $\mu_i$ 's, de modo que  $\mu_{i_1} \leq \mu_{i_2} \leq \dots \leq \mu_{i_k}$ . El objetivo es escoger el sistema con la respuesta esperada más pequeña,  $\mu_{i_1}$  (si se deseara obtener la respuesta esperada más grande,  $\mu_{i_k}$ , bastaría con cambiar el signo de los  $X_{ij}$ 's y los  $\mu_i$ 's).

La naturaleza inherentemente aleatoria de las observaciones  $X_{ij}$ 's implica que nunca podemos estar absolutamente seguros de tomar la decisión correcta, aunque si podemos cuantificar la probabilidad de haberlo hecho.

Notemos CS el evento "tomar la decisión correcta". Si  $\mu_{i_1}$  y  $\mu_{i_2}$  están muy próximos, es indiferente, a efectos prácticos, seleccionar el sistema  $i_1$  (el de media  $\mu_{i_1}$ ) o el  $i_2$  (el de media  $\mu_{i_2}$ ), con lo cual, debemos desarrollar un método que nos permita evitar el gran número de repeticiones necesarias para resolver esta diferencia sin importancia.

Teniendo esto en cuenta, la formulación del problema es:

deseamos tomar la decisión correcta, con una probabilidad al menos  $P^*$ ,  $P\{CS\} \geq P^*$ , supuesto que  $\mu_{i_2} - \mu_{i_1} \geq d^*$ , donde la probabilidad mínima de CS,  $P^* > \frac{1}{k}$ , y la "diferencia despreciable"  $d^*$ , son ambas especificadas por el analista. Es natural preguntarse qué sucede si  $\mu_{i_2} - \mu_{i_1} < d^*$ .

El procedimiento, que explicaremos a continuación, tiene la propiedad deseable de que, con probabilidad, al menos  $P^*$ , el valor verdadero de la media del sistema escogido no será mayor que  $\mu_{i_1} + d^*$ . Así pues, en cualquier caso, no se seleccionará un sistema cuya media sea mayor que la del sistema verdadero en más de  $d^*$  unidades.

El procedimiento, que supone que las  $X_{ij}$ 's están distribuidas (al menos aproximadamente) normalmente, consta de dos etapas de recogida de observaciones de cada uno de los k sistemas.

En la primera etapa se realiza un número fijo de repeticiones independientes de la simulación de cada sistema. A partir de las estimaciones de las varianzas de este primer conjunto de observaciones, se determina el número adicional de replicaciones de la simulación de cada sistema necesario para poder tomar una decisión.

En la primera etapa, se realizan  $n_0 > 2$  replicaciones de la simulación de cada uno de los k sistemas y se definen:

$$\bar{X}_i^{(1)} = \frac{\sum_{j=1}^{n_0} X_{ij}}{n_0}; \quad s_i^2 = \frac{\sum_{j=1}^{n_0} [X_{ij} - \bar{X}_i^{(1)}]^2}{n_0 - 1} \quad \text{para } i=1,2,\dots,k$$

A continuación, se determina el tamaño total de la muestra,  $N_i$ , para el sistema i-ésimo:

$$N_i = \max \left\{ n_0 + 1, \left\lceil \frac{h_1^2 s_i^2}{(d^*)^2} \right\rceil \right\}$$

donde:

$\lceil z \rceil$  es el entero más pequeño, mayor o igual que el real z.

$h_1$  es una constante, dependiente de k,  $P^*$  y  $n_0$ , que puede obtenerse de la siguiente tabla.

$P^*$	$n_0$	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0.90	20	1.896	2.342	2.583	2.747	2.870	2.969	3.051	3.121	3.182
0.90	40	1.852	2.283	2.514	2.669	2.785	2.878	2.954	3.019	3.076
0.95	20	2.453	2.872	3.101	3.258	3.377	3.472	3.551	3.619	3.679
0.95	40	2.386	2.786	3.001	3.150	3.260	3.349	3.422	3.484	3.539

A continuación, se realizan las  $N_i - n_0$  replicaciones adicionales de la simulación del sistema i-ésimo ( $i=1,2,\dots,k$ ) y se calculan las medias de las observaciones obtenidas en esta segunda etapa de simulaciones:

$$\bar{X}_i^{(2)} = \frac{\sum_{j=n_0+1}^{N_i} X_{ij}}{N_i - n_0}$$

Entonces, se definen los pesos:

$$W_{i1} = \frac{n_0}{N_i} \left( 1 + \sqrt{1 - \frac{N_i}{n_0} \left[ 1 - \frac{(N_i - n_0)(d^*)^2}{h_1^2 s_i^2} \right]} \right) \quad \text{para } i=1,2,\dots,k$$

$$W_{i2} = 1 - W_{i1}$$

y las medias ponderadas:

$$\tilde{X}_i = W_{i1} \bar{X}_i^{(1)} + W_{i2} \bar{X}_i^{(2)}$$

Se selecciona el sistema con la  $\tilde{X}_i$  menor.

La elección de  $P^*$  y  $d^*$ , que puede hacerse después de la primera etapa, depende de la precisión necesaria y de la carga computacional que suponga obtener una  $N_i$  grande, asociada con una  $P^*$  grande y una  $d^*$  pequeña. Sin embargo, la elección de  $n_0$  es más problemática: normalmente se recomienda escoger  $n_0$  igual a 20.

**Ejemplo.** Se desean comparar  $k=5$  políticas alternativas de gestión del inventario en base a sus correspondientes estimaciones del coste medio mensual para los primeros 120 meses de operación, que notamos  $\mu_i$  para la política  $i$ -ésima. El objetivo es seleccionar el sistema con menor  $\mu_i$ , con un  $100P^* \% = 90\%$  de probabilidad de hacer la elección correcta, supuesto que  $\mu_{i_2} - \mu_{i_1} \geq d^* = 1$ . Se realizan  $n_0 = 20$  replicaciones independientes de la simulación de cada sistema. De la tabla se obtiene  $h_1 = 2.747$ .

Los resultados de la primera etapa son:

Política $i$ , $(s_i, S_i)$	$\bar{X}_i^{(1)}$	$s_i^2$	$N_i$
$i=1$ , (20,40)	126.484	14.521	110
$i=2$ , (20,80)	121.915	7.961	61
$i=3$ , (40,60)	127.160	9.451	72
$i=4$ , (40,100)	130.710	8.246	63
$i=5$ , (60,100)	144.072	6.197	47

A partir de los  $s_i^2$ ,  $s$ ,  $h_1$ , y  $d^*$ , se ha calculado el valor de total de la muestra,  $N_i$ , para cada sistema.

Así pues, realizamos  $N_i - 20$  replicaciones adicionales de cada política: 90 para la política uno, 41 para la dos, etc... y se calculan las medias de los nuevos resultados obtenidos,  $\bar{X}_i^{(2)}$ . Finalmente, se calculan los pesos  $W_{i1}$  y  $W_{i2}$  para cada sistema, así como la media ponderada,  $\tilde{X}_i$ . Se obtiene:

Política i, $(s_i, S_i)$	$\bar{X}_i^{(2)}$	$W_{i1}$	$W_{i2}$	$\tilde{X}_i$
i=1 , (20,40)	124.453	0.206	0.794	124.871
i=2 , (20,80)	121.633	0.386	0.614	121.742
i=3 , (40,60)	126.105	0.321	0.679	126.444
i=4 , (40,100)	132.031	0.369	0.631	131.543
i=5 , (60,100)	144.833	0.460	0.540	144.483

Dado que  $\tilde{X}_2$  es la menor de las medias ponderadas, la política 2, (s=20,S=80), parece ser la de menor costo.