ANÁLISIS DE LOS DATOS DE SALIDA DE UNA SIMULACIÓN

APUNTE COMPLEMENTARIO

1. ESTIMACIÓN DEL ESTIMADOR DE PUNTO E INTERVALO DE CONFIANZA EN SIMULACIONES TERMINANTES DE VARIAS CORRIDAS

Muestras de la Variable de Salida θ	***Corrida 1***	***Corrida 2***	**…**	***Corrida R***
Observación 1	Y_{11}	Y_{21}	…	Y_{R1}
Observación 2	Y_{12}	Y_{22}	…	Y_{R2}
…	…	…	…	…
Observación n-ésima	Y_{1n_1}	Y_{1n_2}	…	Y_{Rn_r}
Estimador de Punto de la corrida	$\hat{\theta}_1$	$\hat{\theta}_2$	…	$\hat{\theta}_R$

El estimador de punto para cada corrida completa se calcula como:

$$\hat{\theta}_r = \frac{1}{n_r}\sum_{i=1}^{n_r} Y_{ri}, \quad r = 1, \ldots, R$$

donde n_r es la cantidad de observaciones que se efectuaron en la corrida r-ésima.

El estimador de punto general de la variable de salida se calcula como:

$$\hat{\theta} = \frac{1}{R}\sum_{r=1}^{R} \hat{\theta}_r$$

La varianza estimada del estimador de punto de la variable de salida se calcula como:

$$\hat{\sigma}^2(\hat{\theta}) = \frac{1}{(R-1)R}\sum_{r=1}^{R} (\hat{\theta}_r - \hat{\theta})^2$$

Los intervalos de confianza para la variable de salida se calculan como:

$$\hat{\theta} - t_{\alpha/2,f}\hat{\sigma}(\hat{\theta}) \le \theta \le \hat{\theta} + t_{\alpha/2,f}\hat{\sigma}(\hat{\theta})$$

donde:

t es la distribución t-student,
1-α es el nivel de certeza, y
f son los grados de libertad, siendo f = R-1.

Si la cantidad de corridas fuera superior a 30, utilizamos la distribución normal en vez de la t-student debido a que las tablas en general llegan hasta 30 grados de libertad. En tales casos, los intervalos de confianza se computan como:

$$\hat{\theta} - z_{1-\alpha/2}\hat{\sigma}(\hat{\theta}) \le \theta \le \hat{\theta} + z_{1-\alpha/2}\hat{\sigma}(\hat{\theta})$$

2. Estimación del Estimador de Punto e Intervalo de confianza en Simulaciones Terminantes de una Corrida

Muestras de la Variable de Salida θ	***Única Corrida***
Observación 1	Y_1
Observación 2	Y_2
…	…
Observación n	Y_n

En este caso, el estimador de punto general de la variable de salida se calcula directamente como la media de las observaciones:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

La varianza estimada del estimador de punto de la variable de salida se calcula como:

$$\hat{\sigma}^2(\hat{\theta}) = \frac{S^2}{n}$$

donde S^2 es la varianza de la muestra:

$$S^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(Y_i - \hat{\theta})^2$$

Luego, los intervalos de confianza para la variable de salida se calculan igual que antes:

$$\hat{\theta} - t_{\alpha/2,f}\hat{\sigma}(\hat{\theta}) \le \theta \le \hat{\theta} + t_{\alpha/2,f}\hat{\sigma}(\hat{\theta})$$

donde:

t es la distribución t-student,

1-α es el nivel de certeza,

f son los grados de libertad, siendo f = n-1, y

$$\hat{\sigma}(\hat{\theta}) = \frac{S}{\sqrt{n}}$$

Si la cantidad de muestras, es decir, el número n de observaciones de la corrida, fuera superior a 30, utilizamos la distribución normal en vez de la t-student debido a que las tablas en general llegan hasta 30 grados de libertad. En tales casos, los intervalos de confianza se computan como:

$$\hat{\theta} - z_{1-\alpha/2}\hat{\sigma}(\hat{\theta}) \le \theta \le \hat{\theta} + z_{1-\alpha/2}\hat{\sigma}(\hat{\theta})$$

3. COMPARACIÓN ENTRE SIMULACIONES TERMINANTES DE VARIAS CORRIDAS CON MUESTRAS INDEPENDIENTES.

Cuando deseamos comparar los valores obtenidos por dos simulaciones de varias corridas para una misma variable de salida θ con muestras independientes, se dispone de la siguiente información:

SIMULACIÓN DEL SISTEMA 1

Muestras de la Variable de Salida θ	***Corrida 1 del Sistema 1***	***Corrida 2 del Sistema 1***	**…**	***Corrida R1 del Sistema 1***
Observación 1	Y_{11}	Y_{21}	…	$Y_{R_1 1}$
Observación 2	Y_{12}	Y_{22}	…	$Y_{R_1 2}$
…	…	…	…	…
Observación n-ésima	Y_{1n_1}	Y_{1n_2}	…	$Y_{R_1 n_r}$
Estimador de Punto de la corrida	$\hat{\theta}_{S_1 1}$	$\hat{\theta}_{S_1 2}$	…	$\hat{\theta}_{S_1 R_1}$

SIMULACIÓN DEL SISTEMA 2

Muestras de la Variable de Salida θ	***Corrida 1 del Sistema 2***	***Corrida 2 del Sistema 2***	**…**	***Corrida*** R_2 ***del Sistema 2***
Observación 1	Y_{11}	Y_{21}	…	$Y_{R_2 1}$
Observación 2	Y_{12}	Y_{22}	…	$Y_{R_2 2}$
…	…	…	…	…
Observación n-ésima	Y_{1n_1}	Y_{1n_2}	…	$Y_{R_2 n_r}$
Estimador de Punto de la corrida	$\hat{\theta}_{S_2 1}$	$\hat{\theta}_{S_2 2}$	…	$\hat{\theta}_{S_2 R_2}$

El estimador de punto para cada corrida completa del Sistema 1 se calcula como:

$$\hat{\theta}_{S_1 r} = \frac{1}{n_r}\sum_{i=1}^{n_r} Y_{ri}, \quad r = 1,\ldots,R_1$$

donde n_r es la cantidad de observaciones que se efectuaron en la corrida r-ésima.

El estimador de punto general de la variable de salida para el Sistema 1 se calcula como:

$$\hat{\theta}_{S_1} = \frac{1}{R_1}\sum_{r=1}^{R_1} \hat{\theta}_{S_1 r}$$

La varianza estimada del estimador de punto de la variable de salida en el Sistema 1 se calcula como:

$$\hat{\sigma}^2(\hat{\theta}_{S_1}) = \frac{1}{(R_1 - 1)R_1}\sum_{r=1}^{R_1}(\hat{\theta}_{S_1 r} - \hat{\theta}_{S_1})^2$$

Se repite el mismo análisis para el Sistema 2 obteniéndose:

$$\hat{\theta}_{S_2} \text{ y } \hat{\sigma}(\hat{\theta}_{S_2})$$

Luego, las salidas de ambos sistemas para la misma variable de salida se comparan construyendo el siguiente intervalo de confianza:

$$(\hat{\theta}_{S_1} - \hat{\theta}_{S_2} - z_{1-\alpha/2}\sqrt{\hat{\sigma}^2(\hat{\theta}_{S_1}) + \hat{\sigma}^2(\hat{\theta}_{S_2})},\quad \hat{\theta}_{S_1} - \hat{\theta}_{S_2} + z_{1-\alpha/2}\sqrt{\hat{\sigma}^2(\hat{\theta}_{S_1}) + \hat{\sigma}^2(\hat{\theta}_{S_2})})$$

donde:

z corresponde a la distribución normal, y

1-α es el nivel de certeza.

En este caso no podemos emplear la t-student para construir el intervalo por una sencilla razón: si $R_1 \neq R_2$ luego no podemos determinar los grados de libertad.

Una vez construido el intervalo, si el valor 0 está incluido dentro del mismo, se considera despreciable la diferencia obtenida por las simulaciones para la variable de salida θ.

4. COMPARACIÓN ENTRE SIMULACIONES TERMINANTES DE UNA ÚNICA CORRIDA CON MUESTRAS INDEPENDIENTES

Cuando deseamos comparar los valores obtenidos por dos simulaciones de una única corrida para una misma variable de salida θ con muestras independientes, se dispone de la siguiente información:

Muestras de la Variable de Salida θ	*Única Corrida del Sistema 1*	*Única Corrida del Sistema 2*
Observación 1	$Y_{S_1 1}$	$Y_{S_2 1}$
Observación 2	$Y_{S_1 2}$	$Y_{S_2 2}$
…	…	…
Observación n	$Y_{S_1 n_1}$	$Y_{S_2 n_2}$

En este caso, los estimadores de punto generales de la variable de salida para cada sistema se calculan directamente como las medias de las observaciones:

$$\hat{\theta}_{S_1} = \frac{1}{n_1}\sum_{i=1}^{n_1} Y_{s_1 i} \qquad \hat{\theta}_{S_2} = \frac{1}{n_2}\sum_{i=1}^{n_2} Y_{s_2 i}$$

Las varianzas estimadas de los estimadores de punto de la variable de salida para cada sistema se calculan como:

$$\hat{\sigma}^2(\hat{\theta}_{S_1}) = \frac{S_{S_1}^2}{n_1} \qquad \hat{\sigma}^2(\hat{\theta}_{S_2}) = \frac{S_{S_2}^2}{n_2}$$

donde las varianzas de las muestras se computan como:

$$S_{S_1}^2 = \frac{1}{(n_1 - 1)}\sum_{i=1}^{n_1}(Y_{S_1 i} - \hat{\theta}_{S_1})^2 \qquad S_{S_2}^2 = \frac{1}{(n_2 - 1)}\sum_{i=1}^{n_2}(Y_{S_2 i} - \hat{\theta}_{S_2})^2$$

Luego, las salidas de ambos sistemas para la misma variable de salida se comparan construyendo el siguiente intervalo de confianza:

$$(\hat{\theta}_{S_1} - \hat{\theta}_{S_2} - z_{1-\alpha/2}\sqrt{\hat{\sigma}^2(\hat{\theta}_{S_1}) + \hat{\sigma}^2(\hat{\theta}_{S_2})}, \quad \hat{\theta}_{S_1} - \hat{\theta}_{S_2} + z_{1-\alpha/2}\sqrt{\hat{\sigma}^2(\hat{\theta}_{S_1}) + \hat{\sigma}^2(\hat{\theta}_{S_2})})$$

donde:

> z corresponde a la distribución normal, y
> 1-α es el nivel de certeza.

En este caso no podemos emplear la t-student para construir el intervalo por una sencilla razón: si $n_1 \neq n_2$ luego no podemos determinar los grados de libertad.

Al igual que antes, una vez construido el intervalo, si el valor 0 está incluido dentro del mismo, se considera despreciable la diferencia obtenida por las simulaciones para la variable de salida θ .

5. COMPARACIÓN ENTRE SIMULACIONES TERMINANTES DE VARIAS CORRIDAS CON MUESTRAS CORRELACIONADAS.

Cuando deseamos comparar los valores obtenidos por dos simulaciones de varias corridas para una misma variable de salida θ con muestras correlacionadas, se dispone de la siguiente información:

SIMULACIÓN DEL SISTEMA 1

Muestras de la Variable de Salida θ	***Corrida 1 del Sistema 1***	***Corrida 2 del Sistema 1***	**…**	***Corrida R del Sistema 1***
Observación 1	Y_{11}	Y_{21}	…	Y_{R1}
Observación 2	Y_{12}	Y_{22}	…	Y_{R2}
…	…	…	…	…
Observación n-ésima	Y_{1n_1}	Y_{1n_2}	…	Y_{Rn_r}
Estimador de Punto de la corrida	$\hat{\theta}_{S_1 1}$	$\hat{\theta}_{S_1 2}$	…	$\hat{\theta}_{S_1 R}$

SIMULACIÓN DEL SISTEMA 2

Muestras de la Variable de Salida θ	***Corrida 1 del Sistema 2***	***Corrida 2 del Sistema 2***	**…**	***Corrida R del Sistema 2***
Observación 1	Y_{11}	Y_{21}	…	Y_{R1}
Observación 2	Y_{12}	Y_{22}	…	Y_{R2}
…	…	…	…	…
Observación n-ésima	Y_{1n_1}	Y_{1n_2}	…	Y_{Rn_r}
Estimador de Punto de la corrida	$\hat{\theta}_{S_2 1}$	$\hat{\theta}_{S_2 2}$	…	$\hat{\theta}_{S_2 R}$

El estimador de punto para cada corrida completa para cada sistema se calcula como:

$$\hat{\theta}_{S_1 r} = \frac{1}{n_r}\sum_{i=1}^{n_r} Y_{ri}, \quad r = 1,\dots,R \qquad\qquad \hat{\theta}_{S_2 r} = \frac{1}{n_r}\sum_{i=1}^{n_r} Y_{ri}, \quad r = 1,\dots,R$$

donde n_r es la cantidad de observaciones que se efectuaron en la corrida r-ésima. Note que los valores de n_r y el valor de R son iguales para ambos sistemas.

Luego se construye la siguiente tabla:

Número de Corrida	Estimador de Punto del Sistema 1 para la corrida	Estimador de Punto del Sistema 2 para la corrida	Diferencia entre corridas
1	$\hat{\theta}_{S_1 1}$	$\hat{\theta}_{S_2 1}$	$\hat{\theta}_{S_1 1} - \hat{\theta}_{S_2 1}$
2	$\hat{\theta}_{S_1 2}$	$\hat{\theta}_{S_2 2}$	$\hat{\theta}_{S_1 2} - \hat{\theta}_{S_2 2}$
…	…	…	…
R	$\hat{\theta}_{S_1 R}$	$\hat{\theta}_{S_2 R}$	$\hat{\theta}_{S_1 R} - \hat{\theta}_{S_2 R}$

Luego calculamos el estimador de punto de las diferencias como:

$$\hat{D} = \frac{1}{R}\sum_{r=1}^{R}(\hat{\theta}_{S_1 r} - \hat{\theta}_{S_2 r})$$

Después se calcula la varianza estimada del estimador de punto de la diferencia como:

$$\sigma^2(\hat{D}) = \frac{1}{(R-1)R}\sum_{r=1}^{R}((\hat{\theta}_{S_1 r} - \hat{\theta}_{S_2 r}) - \hat{D})^2$$

Finalmente, se construye el siguiente intervalo de confianza:

$$(\hat{D} - t_{\alpha/2,f}\hat{\sigma}(\hat{D}), \hat{D} + t_{\alpha/2,f}\hat{\sigma}(\hat{D}))$$

donde:

t es la distribución t-student,
1-α es el nivel de certeza, y
f son los grados de libertad, siendo f = R-1.

Si la cantidad de corridas fuera superior a 30, utilizamos la distribución normal en vez de la t-student debido a que las tablas en general llegan hasta 30 grados de libertad. En tales casos, los intervalos de confianza se computan como:

$$(\hat{D} - z_{1-\alpha/2}\hat{\sigma}(\hat{D}), \hat{D} + z_{1-\alpha/2}\hat{\sigma}(\hat{D}))$$

Una vez construido el intervalo, si el valor 0 está incluido dentro del mismo, se considera despreciable la diferencia obtenida por las simulaciones para la variable de salida θ .

6. Comparación entre Simulaciones Terminantes de una Única Corrida con Muestras Correlacionadas

Cuando deseamos comparar los valores obtenidos por dos simulaciones de una única corrida para una misma variable de salida θ con muestras correlacionadas, se dispone de la siguiente información:

Muestras de la Variable de Salida θ	***Única Corrida Del Sistema 1***	***Única Corrida Del Sistema 2***
Observación 1	$Y_{S_1 1}$	$Y_{S_2 1}$
Observación 2	$Y_{S_1 2}$	$Y_{S_2 2}$
…	…	…
Observación n	$Y_{S_1 n}$	$Y_{S_2 n}$

El primer paso consiste en construir la tabla de las diferencias:

Muestras de la Variable de Salida θ	***Única Corrida del Sistema 1***	***Única Corrida Del Sistema 2***	***Diferencias entre Sistemas***
Observación 1	$Y_{S_1 1}$	$Y_{S_2 1}$	$Y_{S_1 1} - Y_{S_2 1}$
Observación 2	$Y_{S_1 2}$	$Y_{S_2 2}$	$Y_{S_1 2} - Y_{S_2 2}$
…	…	…	…
Observación n	$Y_{S_1 n}$	$Y_{S_2 n}$	$Y_{S_1 n} - Y_{S_2 n}$

Luego calculamos el estimador de punto de las diferencias como:

$$\hat{D} = \frac{1}{n}\sum_{i=1}^{n}(Y_{S_1 i} - Y_{S_2 i})$$

Después se calcula la varianza estimada del estimador de punto de la diferencia como:

$$\sigma^2(\hat{D}) = \frac{1}{(n-1)n}\sum_{i=1}^{n}((Y_{S_1 i} - Y_{S_2 i}) - \hat{D})^2$$

Finalmente, se construye el siguiente intervalo de confianza:

$$(\hat{D} - t_{\alpha/2,f}\hat{\sigma}(\hat{D}),\ \hat{D} + t_{\alpha/2,f}\hat{\sigma}(\hat{D}))$$

donde:

t es la distribución t-student,
1-α es el nivel de certeza, y
f son los grados de libertad, siendo f = n-1.

Si la cantidad de observaciones fuera superior a 30, utilizamos la distribución normal en vez de la t-student debido a que las tablas en general llegan hasta 30 grados de libertad. En tales casos, los intervalos de confianza se computan como:

$$(\hat{D} - z_{1-\alpha/2}\hat{\sigma}(\hat{D}),\ \hat{D} + z_{1-\alpha/2}\hat{\sigma}(\hat{D}))$$

Una vez construido el intervalo, si el valor 0 está incluido dentro del mismo, se considera despreciable la diferencia obtenida por las simulaciones para la variable de salida θ.