Tarea 5: Introducción a Ciencia de Datos.

Rojas Gutiérrez Rodolfo Emmanuel

15 de octubre de 2021

1. Introduction

A lo largo de esta tarea, $M_{n\times m}(\mathbb{R})$ denota el espacio de todas las matrices de dimensión $n \times m$ con coeficientes en los reales, adicionalmente si $A \in M_{n\times m}(\mathbb{R})$ entonces A' denota a la matriz transpuesta de A.

Ejercicio 1:

Solución. Para este ejercicio, se empleo el conjunto de datos siguiente:[1]

MC	VC	LO	NO	SO
77	82	67	67	81
63	78	80	70	81
75	73	71	66	81
55	72	63	70	68
63	63	65	70	63
53	61	72	64	73
⋮	⋮	⋮	⋮	⋮
5	30	44	36	18
12	30	32	35	21
5	26	15	20	20
0	40	21	9	14

Cuadro 1: Datos de puntuaciones de 88 estudiantes en 5 exámenes de 5 materias, 2 de ellos hechos a libro abierto y los restantes a libro cerrado.

Es importante comentar que para $i \in \{1, \ldots, 88\}$, el renglón i de la tabla de datos 1 corresponde a las calificaciones del estudiante i, en cada uno de los 5 exámenes presentados, adicionalmente, se tiene que las columnas MC y VC son las columnas que corresponden a los exámenes que fueron realizados a libro abierto. Ahora, sea $X \in M_{88\times 5}(\mathbb{R})$ aquella matriz de datos cuyas filas y columnas coinciden con las filas y columnas de la tabla de datos 1 y, suponga que cada observación de 5 estudiantes proviene de una misma distribución 5-variada, con matriz de covarianzas Σ la cual se

[1] Los datos al completo, pueden ser consultados en R.

desconoce. Con este contexto, se pretende encontrar con fines de evaluación aquella combinación de pesos $\{\omega_1, \ldots, \omega_5\}$ que máximize la variabilidad en el conjunto de calificaciones, al proyectar las mismas mediante un promedio ponderado de la forma[2]

$$\omega_1 X_{i1} + \ldots + \omega_5 X_{i5}, \text{ con } i \in \{1, \ldots, 88\}, \tag{1}$$

lo cual, permitiría resumir de la mejor forma posible el desempeño de cada estudiante a través de los 5 exámenes y, diferenciarlo del rendimiento de los demás estudiantes. Como puede imaginar, existen infinidad de opciones para elegir los pesos anteriores, no obstante, muchas de ellas no permitan hacer diferencia alguna entre los estudiantes, por ejemplo, uno podría dar todo el peso al examen SO, es decir, uno podría elegir los pesos $(\omega_1, \ldots, \omega_5)$ como $(0, 0, 0, 0, 1)$, no obstante, si a todos los alumnos les fuera muy bien o muy mal en dicho examen no se estaría haciendo distinción alguna entre ellos. Sin embargo, no debemos rompernos la cabeza desarrollando un algoritmo nuevo para determinar los pesos $(\omega_1, \ldots, \omega_5)$ necesarios para llevar a cabo nuestro cometido, pues, por lo visto en el tema de componentes principales, se sabe que el conjunto de pesos $(\omega_1, \ldots, \omega_5)$ que maximiza la variabilidad en el conjunto de datos proyectados mediante el promedio ponderado 1, puede elegirse como un vector propio $(\omega_1, \ldots, \omega_5)$ de Σ con norma 1 y asociado a su mayor valor propio, sin embargo, dado que se desconoce el valor de Σ, se debe emplear en su lugar a la matriz de covarianzas muestral

$$S = \frac{X_c' X_c}{88}, \tag{2}$$

donde, X_c representa a la matriz de datos X centrada. Este detalle hace una diferencia importante, pues, al ser S una estimación de Σ entonces S depende de los valores de la muestra X, es decir, si hubiésemos observado a otros 88 alumnos presentar los exámenes y, sus calificaciones se distribuyeran de la misma manera que la de los 88 alumnos observados, muy posiblemente obtendríamos una muestra distinta a la presentada en la tabla de datos 1 y, por ende, también se debería obtener un valor de S distinto, de este modo, al calcular los pesos $(\omega_1, \ldots, \omega_5)$ como un vector propio de S con norma 1 asociado al mayor valor propio de S, se obtiene un vector de pesos $\hat{\omega}_1, \ldots, \hat{\omega}_5$ el cual depende del valor de la muestra y, por lo tanto, para cada $i \in \{1, \ldots, 5\}$ la cantidad $\hat{\omega}_i$ tiene asociada a ella cierta variabilidad. Ahora, una pregunta natural es: ¿Cómo podemos estimar esta variabilidad en cada una de las componentes del vector $\hat{\omega}_1, \ldots, \hat{\omega}_5$? La respuesta, mediante remuestreo Bootstrap aplicando el siguiente algoritmo:

1. Para $n \in \{1, \ldots, 10000\}$ se elige una muestra con reemplazo de las 88 filas de X de tamaño 88, con lo que, se forma una matriz $X^{(n)} \in M_{88\times 5}(\mathbb{R})$ cuyas filas son precisamente las filas obtenidas en el muestreo previo.

2. Se calcula la matriz de covarianzas muestral para los datos en la matriz $X^{(n)}$, como:

$$S^{(n)} = \frac{X_c^{(n)\prime} X_c^{(n)}}{88},$$

donde, $X_c^{(n)}$ representa a la matriz $X^{(n)}$ con sus datos centrados. Posteriormente, se obtiene $(\hat{\omega}_1^{(n)}, \ldots, \hat{\omega}_5^{(n)})$, como un vector propio de $X^{(n)}$ con norma 1 asociado al mayor valor propio de $S^{(n)}$.

[2]En el artículo se menciona que se consideran solamente pesos escalados, tales que la suma de los cuadrados de los mismos, sea igual a uno.

	Estimación Puntual	Estimación de la desv. estándar	Valor estimado usando la matriz de covarianzas S de X.
ω_1	0.500	0.0560	0.5054
ω_2	0.3665	0.0416	0.3683
ω_3	0.3439	0.0290	0.3457
ω_4	0.4518	0.0393	0.4511
ω_5	0.5328	0.0454	0.5347

Cuadro 2: Estimaciones puntuales y de variabilidad de los pesos $(\omega_1, \ldots, \omega_5)$, que generan la mayor variabilidad posible en el conjunto de datos proyectados.

3. Se almacenan los valores $(\hat{\omega}_1^{(n)}, \ldots, \hat{\omega}_5^{(n)})$ como la n-ésima columna de una matriz $FC \in M_{5\times 10000}(\mathbb{R})$. Finalmente, si $n = 10000$ el algoritmo acaba y obtenemos una matriz $FC \in M_{5\times 10000}(\mathbb{R})$, cuya i-ésima fila consiste en 10000 estimaciones bootstrap del peso ω_i, con $i \in \{1, \ldots, 5\}$. En otro caso, regrese al paso *1.*

Ahora, el algoritmo anterior posee un problema, dado que los vectores propios calculados en cada iteración no son únicos, pues, al multiplicar un vector propio de norma 1 por menos uno, se obtiene un vector propio con norma 1 y asociado al mismo valor propio, entonces, se adoptó el criterio de considerar únicamente aquellos muestreos en los cuales la primer coordenada de pesos estimados sea positiva, esto con el fin de poder reproducir los resultados en Diaconis y Efron (1983).[3] Así, la matriz FC que resultó tras utilizar el algoritmo en 1, en realidad tiene dimensión 5×9534. Por ende, dado $i \in \{1, \ldots, 5\}$ se tiene que la fila i-ésima de FC, la cual se denotara por FC_i, consiste en 9534 estimaciones Bootstrap para el peso ω_i necesario para maximizar la variabilidad en los datos proyectados (1), por lo que, una forma de obtener un estimador de la variabilidad de la estimación de este peso, es la desviación estándar muestral de FC_i y una estimación puntual de este peso estaría dada por el promedio de los datos en FC_i, bajo esta idea se construyo Tabla presentada en el Cuadro 2: Cabe destacar que, la última columna en la Tabla presentada en el cuadro 2, corresponde a la estimación que se habría obtenido al calcular los pesos deseados, como un valor propio de S con norma 1 asociado al valor propio más grande de S,[4] de este modo, al comparar las columnas uno y tres de la tabla 2 puede ver que dichas estimaciones, y las estimaciones Bootstrap son muy similares. Por último, haciendo uso de los datos de la Tabla (2), se replico una de las gráficas presentadas en Diaconis y Efron (1983), la cual se presenta en la Figura 1. En esta gráfica se puede observar lo que ya se vislumbraba en la segunda columna de la Tabla 2, es decir, que la variabilidad de las estimaciones bootstrap para el vector de pesos es baja. Finalmente, en Diaconis y Efron (1983), se menciona que la segunda componente principal de Σ es el vector de pesos $(w_1, \ldots, w_5)$ que, sujeto a la restricción de independencia, genera la segunda mayor diferencia entre los alumnos, al proyectar con estos pesos los datos de calificaciones como en (1). Sin embargo, para encontrar estos pesos nuevamente nos encontramos ante el problema de que la matriz Σ es desconocida y, por ende, no podemos calcular esta segunda componente simplemente obteniendo un vector propio normalizado de Σ, correspondiente a su segundo valor propio más grande, no obstante, para estos pesos es posible realizar un análisis análogo al realizado para la primer componente de la siguiente forma. Primero, una primer estimación para estos pesos puede obtenerse al calcular el vector propio de S con norma 1 asociado al segundo valor propio más grande de S, ahora, sabemos que esta estimación

[3] Además, de haber sido una recomendación del Dr. Quiroga.

[4] Con S como en (2).

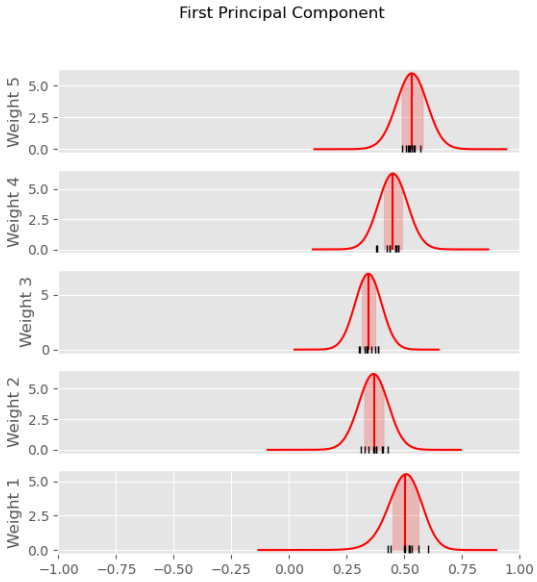

Figura 1: En linea roja solida se presentan las densidades por kernel de las 5 filas de FC, en otras palabras, las densidades por kernel de las estimaciones bootstrap de cada uno de los 5 pesos, igualmente, en lineas rojas solidas verticales puede ver los valores correspondientes de la tercer columna, de la Tabla 2, para cada uno de los pesos $(\omega_1, \ldots, \omega_5)$, por otro lado, el área sombreada en color rojo representa la porción del área bajo la curva que se encuentra a más menos una desviación estándar, segunda columna de la Tabla 2, del valor marcado en linea roja, finalmente, en marcas negras se presentan algunas de las estimaciones bootstrap hechas para cada uno de los pesos.

trae consigo incertidumbre, la cual podemos cuantificar de cierto modo al modificar el algoritmo bootstrap dado para la primer componente principal de la siguiente manera:[5]

1. Para $n \in \{1, \ldots, 10000\}$ se elige una muestra con reemplazo de las 88 filas de X de tamaño 88, con lo que, se forma una matriz $X^{(n)} \in M_{88\times 5}(\mathbb{R})$ cuyas filas son precisamente las filas obtenidas en el muestreo anterior.

2. Se calcula la matriz de covarianzas muestral para los datos en la matriz $X^{(n)}$, como:

$$S^{(n)} = \frac{X_c^{(n)\prime} X_c^{(n)}}{88},$$

donde, $X_c^{(n)}$ representa a la matriz $X^{(n)}$ con sus datos centrados. Posteriormente, se obtiene $(\hat{w}_1^{(n)}, \ldots, \hat{w}_5^{(n)})$, como un vector propio de $X^{(n)}$ con norma 1 asociado al segundo mayor valor propio de $S^{(n)}$.

[5]Con el fin de que los remuestreos de filas, con los que se obtienen las estimaciones bootstrap para la primer y segunda componente principales en cada iteración, sean iguales se debe correr ambos algoritmos al mismo tiempo.

3. Se almacenan los valores $(\hat{w}_1^{(n)}, \ldots, \hat{w}_5^{(n)})$ como la n-ésima columna de una matriz $SC \in M_{5\times 10000}(\mathbb{R})$. Finalmente, si $n = 10000$ el algoritmo acaba y obtenemos una matriz $SC \in M_{5\times 10000}(\mathbb{R})$, cuya fila i-ésima consiste en 10000 estimaciones bootstrap del peso w_i, con $i \in \{1, \ldots, 5\}$. En otro caso, regrese al paso *1.*

Ahora, el algoritmo previo posee el mismo problema que el primer algoritmo dado en 1, por lo que, también se adopto el criterio de considerar únicamente aquellos muestreos en los cuales la primer coordenada de pesos estimados sea positiva, esto con el fin de poder reproducir los resultados en Diaconis y Efron (1983). De este modo, la matriz SC que resultó tras utilizar el algoritmo tiene dimensión 5×7649. Por ende, dado $i \in \{1, \ldots, 5\}$ se tiene que la fila i-ésima de SC, la cual se denotara por SC_i, consiste en 7649 estimaciones Bootstrap para el peso w_i que genera la segunda mayor diferencia variabilidad en los datos proyectados (1), sujeto a la restricción de independencia, por lo que, una forma de obtener un estimador de la variabilidad de la estimación de este peso, es la desviación estándar muestral de SC_i y una segunda estimación puntual de este peso estaría dada por el promedio de los datos en SC_i, con esto en mente, se construyo Tabla presentada en el Cuadro 3:

Pesos	Estimación Puntual	Estimación de la desv. estándar	Valor estimado usando la matriz de covarianzas S de X.
w_1	0.7434	0.0767	0.7487
w_2	0.1918	0.1256	0.2074
w_3	-0.0786	0.0631	-0.0759
w_4	-0.3080	0.1115	-0.3009
w_5	-0.5082	0.1151	-0.5478

Cuadro 3: Estimaciones puntuales y de variabilidad de los pesos $(w_1, \ldots, w_5)$, que generan la segunda mayor variabilidad posible en el conjunto de datos proyectados, sujeto a la restricción de independencia.

Es importante resaltar que la última columna en la Tabla presentada en el cuadro 3, corresponde a la estimación que se habría obtenido al calcular los pesos deseados, como el valor propio de S con norma 1 asociado al segundo valor propio más grande de S, por lo que, se concluye que las dos estimaciones puntuales dadas son similares en los 5 casos.[6] Finalmente, se recreo con los datos de la Tabla 3 la segunda gráfica presentada en Diaconis y Efron (1983), la cual se expone en la Figura 2. De este modo, al comparar las gráficas en 1 y 2 se advierte que, existe mayor variabilidad en los pesos estimados mediante la segunda componente principal, lo cual puede corroborarse al ver las segundas columnas de las Tablas 2 y 3, por último, en el artículo se comenta que el promedio obtenido con los pesos estimados con el segundo componente principal, puede interpretarse como la diferencia entre un promedio de pruebas a libro abierto y un promedio de pruebas a libro cerrado, ahora, los pesos asociados a las pruebas a libro cerrado son $\{w_3, w_4, w_5\}$, mientras que, los pesos asociados a las pruebas a libro cerrado son $\{w_1, w_2\}$,[7] observe que, esta idea se ver reforzada con los resultados en las gráfica presentadas en la Figura 2, pues, las distribuciones bootstrap de los pesos en $\{w_4, w_5\}$ y en $\{w_1, w_2\}$ parecen estar perfectamente separadas, caracterizando a los exámenes a libro cerrado y abierto. Como comentario final, se destaca que la aplicación del remuestreo bootstrap en este

[6] Con S como en (2).

[7] Ver relación en (1) sustituyendo $\omega' s$ por $w' s$.

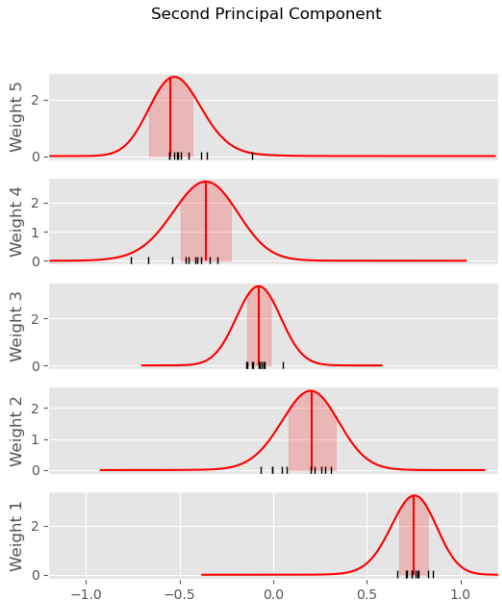

Figura 2: En linea roja solida se presentan las densidades por kernel de las 5 filas de SC, en otras palabras, las densidades por kernel de las estimaciones bootstrap de cada uno de los 5 pesos, igualmente, en lineas rojas solidas verticales puede ver los valores correspondientes de la tercer columna en la Tabla 3 a cada uno de los pesos $(w_1, \ldots, w_5)$, por otro lado, el área sombreada en color rojo representa la porción del área bajo la curva que se encuentra a más menos una desviación estandár, segunda columna de la Tabla 3, del valor marcado en linea roja, finalmente, en marcas negras se presentan algunas de las estimaciones bootstrap hechas para cada uno de los pesos.

problema, es una manera ingeniosa de cuantificar la incertidumbre en las estimaciones dadas, por lo que, se cree que es importante conocer diversas herramientas estadísticas, pues, la combinación de ellas puede resultar en análisis mucho más completos, como en este caso, en el que se ha combinado un análisis de componenetes principales, con una aplicación del remuestreo Bootstrap.

Observación 1. *En el script adjunto, puede encontrar intervalos de confianza al* 95 % *para cada uno de los pesos calculados en el ejercicio, todos ellos construidos con los cuantiles muéstrales de las estimaciones bootstrap. Se comenta esto, pues, por cuestiones de espacio y tiempo no se alcanzo a incluir los mismos en este documento.* △

■

Ejercicio 2:

a)

Solución. Dados n y k dos números naturales tales que $1 \leq k \leq n$, defina $S(n,k)$ como el número de formas de acomodar n objetos en k grupos disjuntos y no vacíos, con la convención de que $S(n-1,n) = 0$, lo cual tiene sentido pues, el número de formas de acomodar $n-1$ objetos en n grupos disjuntos y no vacíos, debe ser cero, dado que no existen suficientes objetos como para formar el número de grupos deseado, adicionalmente, otra convención que será adoptada es que $S(0,0) = 1$. Con este contexto en mente, se argumentará el porque igualdad:

$$S(n,k) = S(n-1,k-1) + kS(n-1,k),$$

es valida para cualesquiera dos números naturales n y k que cumplan la desigualdad $1 \leq k \leq n$. Para ello, sean n y k dos números naturales tales que $1 \leq k \leq n$ y, suponga que tiene n objetos los cuales etiqueta de la siguiente forma $\{o_1, \ldots, o_n\}$, entonces, note que existen únicamente dos maneras de formar k grupos no vacíos y disjuntos a partir de estos n objetos:

1. Seleccionar el primer grupo como $\{o_n\}$, con lo que, resta por acomodar en $k-1$ grupos disjuntos y no vacíos los sobrantes $n-1$ objetos en $\{o_1, \ldots, o_{n-1}\}$, lo cual, por definición de $S(n-1,k-1)$ puede hacerse de exactamente $S(n-1,k-1)$ formas distintas. Así:

   > Existen $S(n-1,k-1)$ formas distintas de acomodar los n objetos en k grupos, cuando uno de los grupos es $\{o_n\}$.

2. Que ninguno de los k grupos sea $\{o_n\}$, con lo que, $\{o_n\}$ debe ser un subconjunto propio de alguno de los k grupos en los que se particione el conjunto $\{o_1, \ldots, o_n\}$,[8] así, si removemos el objeto o_n del grupo al que pertenece, dicho grupo sigue siendo no vacío pues $\{o_n\}$ es un subconjunto propio de tal grupo y, por ende, se obtiene una separación en k grupos disjuntos no vacíos del conjunto $\{o_1, \ldots, o_{n-1}\}$, además, note que si se tiene una partición en k grupos disjuntos no vacíos del conjunto $\{o_1, \ldots, o_{n-1}\}$, entonces, al agregar a cualquiera de estos grupos el elemento o_n, se obtiene una partición del conjunto $\{o_1, \ldots, o_n\}$ en k grupos disjuntos no vacíos. De este modo, el problema de encontrar todas las formas de obtener k grupos disjuntos y no vacíos de $\{o_1, \ldots, o_n\}$, donde, $\{o_n\}$ sea un subconjunto propio de alguno de los k grupos, es equivalente a encontrar todos las formas de acomodar los $n-1$ objetos $\{o_1, \ldots, o_{n-1}\}$ en k grupos disjuntos y no vacios y, posteriormente añadir el objeto o_n a cualquiera de estos k grupos. Ahora, observe que existen $S(n-1,k)$ formas de acomodar los objetos $\{o_1, \ldots, o_{n-1}\}$ en k grupos disjuntos no vacíos y, una vez teniendo una de estas $S(n-1,k)$ agrupaciones, existen k formas de insertar el objeto o_n en alguno de estos k grupos,[9] asi:

   > Existen $kS(n-1,k)$ formas distintas de acomodar los n objetos en k grupos, cuando ninguno de lo k grupos es $\{o_n\}$.

Dado que no existen más posibilidades, se concluye de los dos casos anteriores la formula deseada para las naturales n y k seleccionados, es decir

$$S(n,k) = S(n-1,k-1) + kS(n-1,k). \tag{3}$$

Y, dado que n y k son números naturales tales que $1 \leq k \leq n$ elegidos de forma arbitraria, se concluye el resultado solicitado.

[8] Pues, en otro caso el elemento o_n no se encontraría en ninguno de los k grupos en los que se particiono el conjunto $\{o_1, \ldots, o_n\}$, lo cual sería contradictorio.

[9] Pues, o_n se puede insertar en cualquiera de los k grupos obtenidos.

Observación 2. *Note que, cuando $n = 1$ las convenciones estipuladas son bastante útiles, pues, en este caso solo hay una forma de partir el conjunto $\{o_1\}$ en 1 grupo disjunto y no vació que es precisamente $\{o_1\}$, por lo cual $S(1,1) = 1$ mientras que $S(0,0) = 1$ y $S(0,1) = 0$ por convención, de este modo cuando $n = 1 = k$ se tiene que*

$$S(n-1,k-1) + kS(n-1,k) = S(0,0) + 1 \cdot S(0,1) = 1 = S(1,1) = S(n,k).$$

Ahora, se destaca este caso pues para el mismo el análisis hecho en 1. no es del todo claro. Finalmente, la convención $S(n-1,n) = 0$ es útil cuando $k = n$, pues, en este caso se tiene que existe una única forma de agrupar en n grupos disjuntos no vacíos, a los objetos $\{o_1, \ldots, o_n\}$, la cual consiste en formar un grupo para cada objeto, es decir, particionar $\{o_1, \ldots, o_n\}$ en sus singuletes de la siguiente forma

$$\{o_1, \ldots, o_n\} = \{o_1\} \cup \{o_2\} \cup \ldots \cup \{o_n\},$$

así $S(n,n) = 1$. Finalmente, haciendo un análisis análogo se sigue que $S(n-1,n-1) = 1$ y por convención se tiene que $S(n-1,n) = 0$, por lo cual

$$S(n-1,n-1) + (n-1)S(n-1,n) = 1 + (n-1) * 0 = 1 = S(n,n).$$

Es decir, la formula en (3) en efecto aplica cuando $k = n$. Se destaca la importancia de resaltar este caso, pues, para el mismo el análisis hecho en 2. no es del todo claro. △

■

b)

Solución. La aproximación a emplear es la siguiente: Dado $k \in \mathbb{N}$ se cumple que[10]

$$S(n,k) \approx \frac{k^n}{k!} \text{ cuando } n \to \infty.$$

Lo anterior, es equivalente a

$$\frac{S(n,k)}{k^n/k!} \to 1, \text{ cuando } n \to \infty.$$

Así, dado $k \in \mathbb{N}$ debe existir n suficientemente grande de modo que el cociente entre $S(n,k)$ y $k^n/k!$, sea tan cercano a uno como se desee. En nuestro caso, se tiene $k = 10$ y se supondrá que $n = 1000$ es suficientemente grande como para que el cociente entre $S(n,k)$ y $k^n/k!$ sea cercano a uno, con lo cual, una forma de aproximar $S(n,k)$ es justamente mediante $k^n/k!$, no obstante

$$\frac{k^n}{k!} = \frac{10^n}{10!},$$

es un número tan grande, que incluso el software Julia devolvía infinito cuando se intento evaluar el mismo. Por ende, se decidió emplear un paquete específico en Julia, llamado BigCombinatorics, para calcular $S(n,k)$ cuando $n = 1000$ y $k = 10$. El resultado obtenido aclaro el porque se estaba

[10] Dicha aproximación puede encontrarse dando click aquí y, revisando la sección 5.8 de este artículo de Wikipedia.

obteniendo el resultado de infinito en la aproximación anterior, pues, de acuerdo al paquete citado se tiene que

$$S(n,k) = S(1000,10) = 27557319223985890652557319223985890652557319175819244606408410227624156217900\ldots$$
$$5076870341091064947558923935232396842204745022714288327530129735717170307319095322915997491441133 7$$
$$9982863793843177002018276990645186003191418936645880147240191016770014335843685757340510871136637 3$$
$$3326712187988280471413114695165527579301182484298326671957439022946882788124670680940407335115534 7$$
$$8834713072934899676049862875823552952940263333073887754241815001076806126981915526096049701755444 5$$
$$9927715711462172484389554564451572536659235580694930111124204642360853835932538200761932141101223 6$$
$$1976285829108747907328896777952841254396283166368113062488965007972641703840376938569647263827074 8$$
$$0616456850817094014490605309471229845724849850938291491729443959349489108689748687520002340192718 7$$
$$1730213229219395341285732071618339324933270725103782640004076710437301799410230446974482548630591 9$$
$$1921920705755612206303581672239643950713829187681748502794419033667715959241433673878199302318614 3$$
$$1986569011447354058926234989397588 0.$$

■

Referencias

Diaconis, P. & Efron, B. (1983). Computer-Intensive Methods in Statistics. *Scientific American*, *248*(5), 116-130. https://doi.org/10.1038/scientificamerican0583-116