

Analisis del mercado de aguacates en Estados Unidos

José Pablo Ortiz,
CIMAT

2 de mayo de 2022

1. Introducción

En este trabajo se va analizar el mercado de aguacates en Estados Unidos. En la primera parte de este trabajo se hará un análisis de los precios y montos de aguacates desde una perspectiva que no considere la relación de estas cantidades con el tiempo. Para esto se va a sintetizar la información importante en tablas, se realizaran pruebas de hipótesis para hacer comparaciones multiples, y se propondrá rankings de la relación que guarda el precio con el tamaño, lugar de origen y mes del año. En la segunda parte se presentaran los resultados obtenidos de modelar la serie de tiempo del precio del aguacate utilizando la metodología ARIMA o Box-Jenkins.

2. Múltiples comparaciones

En esta sección se busca encontrar la relación que existe entre el precio del aguacate y su tipo, variedad, tamaño, lugar de origen y mes del año. Para esto se utilizara pruebas de hipótesis. En los casos de tipo y variedad se empleara la prueba Mann-Whitney. Esto se debe a que estas variables solo tienen dos categorías: Hass y *variedades verdes* en el caso de variedad y orgánica y convencional en el caso de tipo. La prueba Mann-Whitney compara dos poblaciones X y Y y tiene como hipótesis nula que la probabilidad de que X sea mayor que Y es igual que la probabilidad de que Y sea mayor X . La hipótesis alternativa es que la probabilidad X sea mayor que Y es menor que la probabilidad de que X sea menor que Y . Así, si se rechaza la hipótesis nula se estaría encontrando evidencia a favor de que Y es mayor que X .

¿La diferencia en precios es estadísticamente diferente para los diferentes tipos de fruta?

El Cuadro 3 muestra las estadísticas descriptivas del tipo de fruta. La experiencia dicta que la fruta orgánica es más cara que la convencional. Es por esto que en este caso se utiliza la hipótesis alternativa de que el precio de la fruta orgánica es mayor que el de la convencional. El p -valor obtenido es de casi cero y, por lo tanto, se acepta este hecho.

Tipo	Mínimo	Cuantil 0.25	Cuantil 0.5	Cuantil 0.75	Máximo	Media
Conventional	5.5	21.5	30.25	41.75	90.25	32.55
Organic	20.25	32.75	42.25	58.75	83.75	45.62

Cuadro 1: Estadísticas descriptivas del tipo de fruta.

¿El precio es estadísticamente diferente para diferentes variedades?

Existen varios tipos de variedades de aguacate: Haas, Arad, Ettinger, fuerte, Pinkerton, Reed, entre otras. En la base de datos de los precios las variedades distintas a la Haas se consideran como una sola categoría: *distintas variedades verdes*. Las estadísticas descriptivas para la variedad Haas y las *distintas variedades verdes* se muestran en el Cuadro 2. Esta información propone que en la prueba Mann Whitney se utilice como hipótesis nula que la variedad HAAS es tiene un costo más elevado que las *distintas variedades verdes*. Al igual que en la pregunta anterior, se obtiene un p -valor cercano a cero y se rechaza la hipótesis nula. Luego se puede concluir que la fruta Hass es más cara que las *variedades verdes*.

Variedad	Minimo	Cuantil 0.25	Cuantil 0.5	Cuantil 0.75	Maximo	Media
Hass	11.25	27.25	36.25	47.75	90.25	38.57
<i>Variedades verdes</i>	5.5	13.0	17.5	24.0	51.5	18.60

Cuadro 2: Estadísticas descriptivas de la variedad de la fruta.

Utilizar este mismo método para las relaciones faltantes no es posible debido a que cada variable tiene más de dos categorías. Este implica que se tienen que hacer uso de técnicas de comparación múltiple. Se utilizaran las pruebas Kruskal-Wallis y la de Dunn con ajuste de Bonferroni para hacer estas comparaciones. La prueba Kruskal-Wallis tiene como hipótesis nula que las diferentes poblaciones tienen una misma distribución, así al rechazarla significaría que sí existe un efecto entre la relación de las diferentes categorías y el precio. Luego, si se rechaza la hipótesis nula de Kruskal-Wallis se utiliza la prueba de Dunn con corrección de Bonferroni para hacer comparaciones entre todas las parejas.

Se tienen dos objetivos en mente. El primero, buscar cuando no existe una diferencia significativa entre dos o más categorías y segundo, hacer un ranking de la relación que existe entre el precio y las diferentes categorías.

¿El precio por tamaño es estadísticamente diferente ?

Para contestar esta pregunta se utilizaran los datos únicamente de la fruta convencional mexicana. Los únicos tamaños de fruta convencional mexicana de los que se tienen datos son de 32, 36, 40, 48, 60, 70 y 84. Al observar el Cuadro 3 se nota que las estadísticas para los tamaños 32, 36, 40 y 48 son muy similares. Por el contrario, las estadísticas de los tamaños 60, 70 y 84 no comparten ninguna similaridad con los demás tamaños.

Tamaño	Minimo	Cuantil 0.25	Cuantil 0.5	Cuantil 0.75	Maximo	Media
32s	16.25	31.75	38.25	48.75	90.25	40.89
36s	16.75	32.25	38.25	48.75	90.25	40.93
40s	17.75	31.75	37.75	49.25	90.25	40.52
48s	20.25	30.75	37.25	48.25	90.25	40.40
60s	17.25	25.75	32.75	41.25	85.25	35.34
70s	15.25	22.875	27.75	34.75	81.75	30.56
84s	11.25	17.75	20.75	24.75	59.25	22.72

Cuadro 3: Estadísticas descriptivas del tamaño de la fruta.

Al realizar la prueba de Kruskal-Wallis se obtiene un p -valor cercano a cero, por lo tanto, se rechaza la hipótesis nula. Luego, se puede utilizar la prueba de Dunn. Los resultados de esta prueba es

que los tamaños 32, 36, 40 y 48 no presentan una diferencia significativa en el precio. Además, los tamaños 60, 70 y 84 son diferentes todos entre si. Luego al tomar la información del Cuadro 3 se hace el siguiente ranqueo de los precios según el tamaño.

Grupo	Tamaños
1	32s, 36s, 40s, 48s
2	60s
3	70s
4	84s

Cuadro 4: A table beside a figure

Estos resultados están en con la practica. La semilla de un aguacate es muy grande, entonces entre más pequeño sea un aguacate(que se de mayor de tamaño) menor cantidad de alimento tendra y más barato será la fruta. Una cuestión interesante de este resultado es que los compradores son indiferente entre los aguacate de tamaño 32s, 36s, 40s, y 48s.

¿Es estadísticamente diferente el precio de la fruta por puerto de origen?

En esta pregunta se utilizarán todos los datos. Los lugares de origen que se consideran para esta pregunta son: México, California, Peru, Chile y Republica Dominicana. El Cuadro 5 muestran las estadísticas relevantes para los lugares de origen. Como dato extra se menciona que México exporta alrededor del 76 % de los aguacates a Estados Unidos.

Tamaño	Minimo	Cuantil 0.25	Cuantil 0.5	Cuantil 0.75	Maximo	Media	10000lbs	%
Mexico	11.25	25.75	34.25	43.25	90.25	36.37	1137760	0.766
Chile	23.25	23.25	23.25	24.0	26.25	24.0	20615	0.013
California	12.25	33.25	42.25	54.25	83.75	43.36	187424	0.126
Peru	17.0	26.125	29.25	33.25	69.0	32.04	81640	
Florida	5.5	10.5	14.5	21.0	51.5	16.26	21434	0.014
RD	13.0	18.0	23.0	30.0	44.5	23.87	34296	0.023

Cuadro 5: A table beside a figure

Se utiliza la misma metodología que en la pregunta anterior. El p -valor de la prueba Kruskal-Wallis es casi igual a 0. Los resultados de la prueba de Dunn nos indica que solo existe un empate -entre Chile y República Dominicana- y al tomar en cuenta el Cuadro 5 se propone el siguiente ranking.

Ranking	Origen
1	California
2	México
3	Peru
3	Chile, República Dominicana
4	Florida

Cuadro 6: Ranking de los precios.

¿El precio promedio por mes es estadísticamente diferente?

Ahora se va a estudiar la relación que tiene el precio con el mes. Para contestar esta pregunta se utilizaran los datos únicamente de la fruta convencional mexicana. Se decidió tomar como

referencia el mes del año en lugar de la semana debido al número de comparaciones que se tienen que hacer. Para comparar los meses del año se tienen que realizar 66 comparaciones; mientras que si se tomaran las semanas del año se tendrían que hacer 1326. El Cuadro 7 muestra la información importante.

Mes	Minimo	Cuantil 0.25	Cuantil 0.5	Cuantil 0.75	Maximo	Media	10000lbs	%
1	10.25	21.25	24.75	30.75	42.25	25.55	150235	0.101
2	13.0	23.25	26.25	34.25	57.75	29.68	117700	0.079
3	12.25	26.25	32.75	44.25	67.75	35.40	130361	0.087
4	11.25	29.75	39.25	49.25	65.25	39.64	129945	0.087
5	13.25	33.25	39.25	48.75	72.75	40.75	126939	0.085
6	8.0	28.75	39.25	48.25	85.25	40.36	134933	0.090
7	5.5	23.5	34.25	46.75	83.75	36.08	134356	0.090
8	5.5	22.0	34.25	52.25	82.75	36.80	122744	0.082
9	7.5	21.0	28.25	41.25	90.25	33.61	118708	0.079
10	8.5	18.0	26.0	33.25	68.75	27.48	122205	0.082
11	8.0	14.5	18.25	27.75	59.75	21.25	90115	0.060
12	8.5	16.75	19.0	27.25	44.75	22.11	106367	0.071

Cuadro 7: Estadísticas descriptivas de los meses del año.

Al utilizar la prueba de Kruskal-Wallis y después la prueba de Dunn se llega al siguiente ranking

Ranking	Origen
1	4,5,6
2	3,7,8
3	9
4	2
5	1,10
6	11,12

Cuadro 8: Ranking de los precios.

3. Serie de Tiempo

En la base de datos de los precios hay dos medidas de los precios *Low Price* y *High Price*. La frecuencia con la que se observan estas medidas es diaria. No obstante, para este análisis se tomaran los siguientes datos semanales: precio mínimo, precio promedio y precio máximo. Se pretende hacer predicciones de estos 3 datos para cada tamaño significativamente diferente, ver Cuadro 4. Las predicciones solo se harán de la fruta convencional mexicana.

En los siguientes Cuadros se muestra el modelo ARIMA obtenido, las predicciones de 4 semanas en el futuro y los datos observados de esas 4 semanas.

Tamaño	Mínimo	Promedio	Máximo
32s, 36s, 40s, 48s	(3,0,3)	(3,0,2)	(3,0,3)
60s	(2,0,3)	(2,0,3)	8(2,0,0)
70s	(2,0,1)	(2,0,1)	(2,0,1)
84s	(2,0,0)	(2,0,1)	(2,0,0j)

Cuadro 9: Ranking de los precios.