

"Análisis de Varianza Funcional para medir el efecto de recubrimientos revestibles de almidón en características físico-químicas de papayas durante las dos primeras semanas del período de maduración post-cosecha"

Autor: Christian Galarza Morales⁽¹⁾ Coautor: Francisco Vera Alcívar⁽²⁾

Instituto de Ciencias Matemáticas
Escuela Superior Politécnica del Litoral
Campos Gustavo Galindo V.
Km 30.5 Vía Perimetral contiguo a Santa Cecilia, 09-01-5863,
Guayaquil, Ecuador
chedgala@espol.edu.ec⁽¹⁾ fvera@espol.edu.ec⁽²⁾

Resumen

El mercado actual demanda productos de alta calidad y de consumo inmediato, que contengan sólo ingredientes naturales. El éxito de estos productos se debe a sus buenas características sensoriales y nutricionales. Por ello, ha sido necesario el desarrollo de nuevas tecnologías de conservación que retrasen los procesos de deterioro y mantengan la calidad e inocuidad microbiológica, durante la comercialización de estos productos. Una alternativa es el uso de películas comestibles biodegradables las cuales crean una barrera a los gases, produciendo una atmósfera modificada alrededor del producto reduciendo la tasa de respiración y la pérdida de agua, aumentando así, la vida de "anaquel".

La orientación de este estudio es realizar un Análisis de Varianza Funcional con el objetivo de determinar la efectividad del uso de dos diferentes tipos de películas comestibles en frutas al preservar el producto durante su período de maduración pos-cosecha. La ventaja de la aplicación de un Análisis de Datos Funcionales (FDA), es explotar las características deseables que tienen las funciones, como el disponer de estimaciones para valores no observados, el poder aplicar toda técnica estadística univariada como multivariada, así como utilizar las propiedades de sus funciones derivadas las cuales son una fuente de información adicional de los aspectos dinámicos del modelo. El lector deberá tener conocimientos previos en Álgebra Lineal, Estadística Multivariada, Simulación Matemática y Diseño de Experimentos.

Palabras Claves: *Análisis de datos funcionales, Análisis Funcional de varianza, FANOVA, Películas comestibles, Diseño Experimental, Papayas*

Abstract

The current market demand for high quality products and immediate consumption, containing only natural ingredients. The success of these products is due to its good nutritional and sensory characteristics. It has therefore been necessary to develop new conservation technologies that delay deterioration processes and maintain microbiological safety and quality during the marketing of these products. An alternative is the use of biodegradable edible films which create a barrier to

gases, producing a modified atmosphere around the product by reducing the respiration rate and the loss of water, thereby increasing the life of "shelf".

The focus of this study is to perform a functional analysis of variance in order to determine the effectiveness of using two different types of edible films in fruit to preserve the product during the period of post-harvest ripening. The advantage of applying a Functional Data Analysis is to exploit the desirable characteristics that have functions such as having estimations for unobserved values, the application of any univariate and multivariate statistical technique, and use the properties of their derivatives which are an additional source of information for the dynamic aspects of the model. The reader must have previous knowledge of Linear Algebra, Multivariate Statistics, Mathematical Simulation and Design of Experiments.

Keywords: *Functional Data Analysis, Functional Analysis of variance, FANOVA, Edible Film, Experimental Design, Papayas*

1. Introducción

El análisis de datos funcionales es una rama de la estadística que analiza los datos provenientes de curvas, superficies, entre otros, cuyo dominio sea continuo. La filosofía básica del análisis de datos funcionales es tomar las funciones observadas como entidades únicas, mas no como una secuencia de observaciones individuales [11].

1.1. Características

La aplicación de análisis de datos funcionales tiene características deseables que no se tendrían si al contrario realizáramos un análisis estadístico de datos individuales. Algunas de las principales características mencionadas por [6] son:

- No se requieren supuestos paramétricos para la modelación.
- Por lo general, los procesos tienen alta afijación y poco ruido, además se puede suponer que el proceso que genera los datos es suave y continuo.
- Las funciones son diferenciables y derivables por lo son una fuente adicional de información acerca de los aspectos dinámicos del modelo.
- Toda técnica estadística, ya sea descriptiva, inferencial o multivariada puede ser aplicada para datos funcionales.

1.2. Generalidades

Muchas de las técnicas estadísticas multivariadas se extienden al análisis de funciones como observaciones tales como análisis de componentes principales, análisis

de conglomerados, correlación canónica, análisis discriminante, modelos lineales en general entre otras [7].

El análisis de datos funcionales hereda muchos de sus conceptos del análisis funcional, una rama del análisis matemático donde una función representa únicamente un punto en un espacio funcional [8].

En algunos casos, el análisis de datos funcionales permitirá al investigador contestar preguntas no muy fáciles de responder, o que serían computacionalmente complicadas de responder usando los métodos estadísticos tradicionales [6]; más aún, los modelos funcionales facilitan la representación visual de los datos, y esto provee un mayor poder de explicación.

1. 2. Análisis de Datos Funcionales

En la práctica, los datos funcionales son n pares ordenados (t_j, y_j) observados de manera discreta donde y_j es un valor de la función $x(t_j)$ en el tiempo t_j , usualmente afectado por un error ϵ_j . Esto es expresado por

$$y_i = x(t_j) + \epsilon_j$$

o simplemente se puede expresarlo de manera algebraica como

$$\mathbf{y} = \mathbf{x}(\mathbf{t}) + \mathbf{e}$$

donde \mathbf{y} , $\mathbf{x}(\mathbf{t})$ y \mathbf{e} son vectores de longitud n .

2.1. Representación de funciones por medio de sistemas bases

Un sistema de función base es un conjunto de funciones conocidas ϕ_k que son linealmente independientes entre sí y tienen la propiedad que pueden

aproximar arbitrariamente sin problema alguno una función por medio de una combinación lineal de un número suficiente de K funciones de estas [R2]. Un sistema de función base representa una función x por una expansión lineal

$$x(t) = \sum_{k=1}^K c_k \phi_k(t).$$

2.1.1. Sistemas bases. La familia de sistemas de función base tradicional es el conjunto de monomios que son usualmente construidas por diversos tipos de bases. Las más comunes son las series de potencia

$$1, t, t^2, t^3, \dots, t^k, \dots$$

y junto a las series de potencia se encuentran los conocidos sistemas de series de Fourier,

$$1, \sin(\omega t), \cos(\omega t), \sin(2\omega t), \cos(2\omega t) \\ \dots \sin(k\omega t), \cos(k\omega t), \dots$$

los cuales son de gran aplicabilidad, en casos donde los datos son de carácter periódico.

Existen otros tipos de bases como son las ondas (*wavelets*), polinómicas, funciones escalón, b-splines, poligonales, constantes, empíricas, exponenciales, entre otros.

2.2. Ajuste de los datos usando un sistema base por mínimos cuadrados ponderados

Para mínimos cuadrados ponderados la suma cuadrática del error estará dada por

$$SSE(\mathbf{y}|\mathbf{c}) = (\mathbf{y} - \Phi\mathbf{c})' \mathbf{W} (\mathbf{y} - \Phi\mathbf{c}),$$

donde la matriz de pesos puede ser estimada por medio de $\hat{\Sigma}_e^{-1}$ y siendo $\hat{\mathbf{c}}$ el vector de coeficientes que minimiza la expresión anterior calculado por

$$\hat{\mathbf{c}} = (\Phi' \mathbf{W} \Phi)^{-1} \Phi' \mathbf{W} \mathbf{y}.$$

El vector $\hat{\mathbf{y}}$ de valores ajustados será

$$\hat{\mathbf{y}} = \Phi \hat{\mathbf{c}} = \Phi (\Phi' \mathbf{W} \Phi)^{-1} \Phi' \mathbf{W} \mathbf{y} = \mathbf{S}_\Phi \mathbf{y}$$

donde \mathbf{S}_Φ es la matriz “hat” u operador de proyección la cual es idempotente.

2.3. Cuantificación de aspereza (Roughness)

Una manera sencilla de cuantificar la aspereza de una función es el cuadrado de la segunda derivada de una función para todo instante t , esto es $[D^2 x(t)]^2$, lo que usualmente es llamada medida de curvatura en el instante t . La aspereza se medirá por

$$PEN_2(x) = \int [D^2 x(t)]^2 dt.$$

Con lo anteriormente mencionado, podríamos generalizar la penalización de aspereza de orden m como

$$PEN_m(x) = \int [D^m x(t)]^2 dt.$$

Partiendo de la expresión anterior, se puede expresar la penalización de aspereza $PEN_m(x)$ como $\mathbf{c}' \mathbf{R} \mathbf{c}$, donde

$$\mathbf{R} = \int D^m \Phi(t) D^m \Phi(t) dt.$$

2.4. Método de Validación Cruzada Generalizada

El método de validación cruzada generalizada o método GCV, esto por su acrónimo en lengua inglesa *generalized cross-validation*, es una de las medidas más frecuentemente mencionadas en la literatura de suavizamiento, el cual tiene como principal objetivo determinar el valor óptimo del estimador del parámetro de suavización λ .

Según [11], el mejor estimador será el valor de λ que minimice la expresión

$$GCV(\lambda) = \left(\frac{n}{n - df(\lambda)} \right) \left(\frac{SSE}{n - df(\lambda)} \right)$$

donde el segundo multiplicando es el estimador insesgado de la varianza del error σ^2 .

2.5. Suma Cuadrática del Error penalizada de orden m

La Suma Cuadrática del Error penalizada de orden m denotada por $PENSSE_m$, se obtiene sumando la Suma Cuadrática del Error y la penalización de orden m ; esta última se debe multiplicar por el parámetro de suavización λ obtenido previamente por el Método de Validación Cruzada Generalizada. De esta manera tenemos

$$PENSSE_m(\mathbf{y}|\mathbf{c}) = SSE(\mathbf{y}|\mathbf{c}) + \lambda PEN_m(\mathbf{x}) \\ PENSSE_m(\mathbf{y}|\mathbf{c}) = (\mathbf{y} - \Phi\mathbf{c})' \mathbf{W} (\mathbf{y} - \Phi\mathbf{c}) \\ + \lambda \mathbf{c}' \mathbf{R} \mathbf{c}.$$

El vector de estimadores $\hat{\mathbf{c}}$ que minimiza la expresión anterior está dada por

$$\hat{\mathbf{c}} = (\Phi' \mathbf{W} \Phi + \lambda \mathbf{R})^{-1} \Phi' \mathbf{W} \mathbf{y}$$

expresión que coincide la Regularización de Tikhonov cuando $\Gamma = \sqrt{\lambda} \mathbf{R}^{1/2}$.

3. Análisis de Varianza Funcional

3.1. Introducción

El Análisis de Varianza (ANOVA) es una de las herramientas más usadas en la estadística aplicada. Mientras ésta es muy útil cuando se trabaja con datos de pequeñas dimensiones, tiene sus limitaciones analizando

variables de interés funcionales.

En tales casos, el método de Análisis de Varianza Funcional (FANOVA) provee igual solución que el tradicional ANOVA cuando la variable dependiente es funcional, sin dejar de lado su poder de explicación y facilidad de interpretación.

3.2. Aplicación del Análisis Funcional de Varianza

En términos formales, si Y es la característica a medir en la papaya y tenemos P tipos de películas comestibles, el modelo que explicaría la r -ésima observación funcional con el tipo de película p denotada por Y_{pr} sería

$$Y_{pr}(t) = \beta_0(t) + \alpha_p(t) + \epsilon_{pr}(t)$$

donde $\beta_0(t)$ es la función promedio de la característica de interés Y para papayas sin película, $\alpha_p(t)$ es el efecto sobre la característica Y al usar el tipo de película p . El término $\epsilon_{pr}(t)$ representa la función residual, la cual es la variación no explicada respecto a la r -ésima papaya con película comestible p . Nótese que es de interés analizar $\alpha_1(t)$ y $\alpha_2(t)$ dada la condición que

$$\alpha_0(t) = 0, \quad \forall t$$

Podemos definir las tres funciones de regresión β_j correspondientes a las constantes del modelo anterior, esto sería $\beta_1 = \alpha_1$ y $\beta_2 = \alpha_2$. De esta manera el modelo puede ser expresado por

$$Y_{pr}(t) = \sum_j Z_{(pr)j} \beta_j(t) + \epsilon_{pr}(t),$$

o bien podríamos expresarlo de manera compacta en forma matricial por medio de la expresión

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

donde \mathbf{Z} es la matriz de diseño y

$$\boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2].$$

3.3. Ajuste del Modelo

Si tuviéremos un modelo lineal estándar se debería usar el método de Mínimos Cuadrados para hallar el vector $\boldsymbol{\beta}$ que minimice la Suma Cuadrática Residual. Extendiendo el criterio de Mínimos Cuadrados de manera adecuada para el caso funcional, tenemos que cada uno de los residuos $Y_{pr}(t) - \mathbf{Z}_{(pr)}\boldsymbol{\beta}(t)$ constituyen una función, por lo que el criterio de ajuste por Mínimos Cuadrados no ponderados, LMSSE($\boldsymbol{\beta}$) sería

$$\sum_p \sum_r \int_{t_0}^{t_f} \left[Y_{pr}(t) - \sum_{j=0}^p Z_{(pr)j} \beta_j(t) \right]^2 dt,$$

donde t_0 y t_f representan el tiempo inicial y final de la realización del experimento respectivamente. Téngase en

cuenta, que la minimización en realidad es respecto a el vector de coeficientes \mathbf{c} , dado que los $\beta_j(t)$ son combinaciones lineales de los elementos del sistema base Φ .

Bajo la restricción que $Z_{(pr),0} = 1$ se obtiene la estimación del vector de parámetros $\boldsymbol{\beta}$ por Mínimos Cuadrados, siendo la medida cuantitativa del error el área bajo la curva de las funciones residuales al cuadrado.

3.4. Validación del modelo

Para la validación del modelo al igual que en el análisis de varianza común utilizamos el Coeficiente de Determinación R^2 y el Estadístico de prueba F de Fisher. La Potencia de Explicación del Modelo puede ser calculada por medio de la función de Correlación Múltiple Cuadrada RSQ análoga del coeficiente R^2 usado en Análisis de Varianza Simple, esta es calculada como

$$RSQ(t) = \frac{[SST(t) - SSE(t)]}{SST(t)}.$$

Tabla 1: Tabla FANOVA.

Fuentes de Variación	GL	Sumas Cuadráticas	Medias Cuadráticas
REGRESION	$p - 1$	$\sum_{rp} [Z_{pr}\hat{\boldsymbol{\beta}}(t) - \hat{\mu}(t)]^2$	$\frac{SSR(t)}{df_{(regresión)}}$
ERROR (Residuales)	$n - p$	$\sum_{rp} [Y_{pr}(t) - Z_{pr}\hat{\boldsymbol{\beta}}(t)]^2$	$\frac{SSE(t)}{df_{(error)}}$
TOTAL	$n - 1$	$\sum_{rp} [Y_{pr}(t) - \hat{\mu}(t)]^2$	

La función análoga al estadístico F de Fisher llamado en literatura funcional como F_{RATIO} la cual se calcula como

$$F_{RATIO}(t) = \frac{MSR(t)}{MSE(t)}$$

y con esto hallar la función P_{VALUE} equivalente al valor p correspondiente a la Prueba F , la cual nos permitirá concluir si en algún instante de tiempo t durante el experimento, existe algún efecto significativo sobre la variable de interés Y al aplicar un cierto tipo de película. La función P_{VALUE} es calculada como

$$P\left(F_{(df_{(reg)}, df_{(error)})} \geq F_{RATIO}(t)\right).$$

4. Imputación de Datos

4.1. Justificación

Existen características de interés que requieren que el fruto sea cortado para poder ser medidas, es por esto que r papayas seleccionadas aleatoriamente deberán ser cortadas cada día con este objetivo, por lo que se tendrán mediciones faltantes para características que son simplemente observables, por lo que será necesario utilizar un método de imputación de datos para poder estimar estos valores.

4.2. Modelo de Imputación

Básicamente, las observaciones faltantes O_i serán calculadas utilizando la función media \bar{y}_i , de tal manera que restándole una función lineal f_i cuyos coeficientes β_0 y β_1 son de desconocidos y de interés, se ajuste a los datos observados de mejor manera por medio del criterio de Mínimos Cuadrados, esto es

$$O_i = \bar{y}_i - f_i + \epsilon_i$$

siendo f_i una función lineal de la forma

$$f_i = \beta_0 + \beta_1 t_i,$$

ϵ_i es el error, el cual sigue los mismos supuestos que en Regresión Lineal. La Suma Cuadrática del Error estará dada por la expresión

$$Q = \sum_{i=1}^t [\epsilon_i^2] = \sum_{i=1}^t [O_i - \bar{y} + f_i]^2,$$

donde, si llamamos

$$Y_i = (\bar{y} - O_i),$$

y reemplazando f_i la expresión resultante será

$$\sum_{i=1}^t [Y_i - (\beta_0 + \beta_1 t_i)]^2,$$

siendo la expresión de la Suma Cuadrática del Error a minimizar por el Método de Regresión Lineal Simple [16].

5. Experimentación

El principal objetivo del experimento es medir si el uso de películas comestibles en papayas puede retrasar significativamente el proceso de madurez. Por esto, el análisis se centra en ciertas características físico químicas de interés para el consumidor, y que pueden ser indicadores de la maduración del producto.

5.1. Variables

5.1.1. Variable independiente. El factor a considerarse fue el tipo de película comestible a aplicarse. Esta variable categórica toma tres posibles valores que son:

- *Placebo o control.*
- *Película comestible estándar*

- *Película comestible con aceites esenciales:* Película comestible estándar con aceites esenciales de canela y clavo de olor, con el fin de controlar de las pudriciones fungosas.

5.1.2. Variables de Interés. Fueron de interés las características que son percibidas por el consumidor como “sinónimo” de calidad y que al mismo tiempo, reflejan la etapa de maduración y próximo perecimiento de una fruta. Las variables de interés son: peso (gramos), nivel de acidez (pH), sólidos solubles (°brix), dureza (g/cm²) y color, cuya medida de intensidad, fue el módulo del vector formado por sus coordenadas RGB, esto es $\|V\| = \sqrt{R^2 + G^2 + B^2}$.

Cada día se cortarían r papayas de cada nivel para realizar los analizar la acidez pH, los sólidos solubles y dureza, por lo que estas tres características contarán con r observaciones a lo largo del experimento, necesiándose un total de $r * t$ papayas por nivel al inicio del experimento. Las mediciones aunque son tomadas de papayas diferentes, serán relacionadas entre sí, por el nivel de repisa vertical en la refrigeradora donde se coloquen las mismas. El efecto que pudiese tener la posición espacial dentro de las refrigeradoras será confundido en bloques.

Inicialmente, se seleccionarán r papayas de manera aleatoria, las cuales serán las últimas en ser cortadas, y estas serán objeto del análisis de color. Dado que las papayas son cortadas, para la variable peso se utilizó el método de imputación definido en el Capítulo 3, de tal manera que se cuente con la información de las $r * t$ papayas por nivel para todo instante de tiempo t .

5.2. Experimentación Piloto

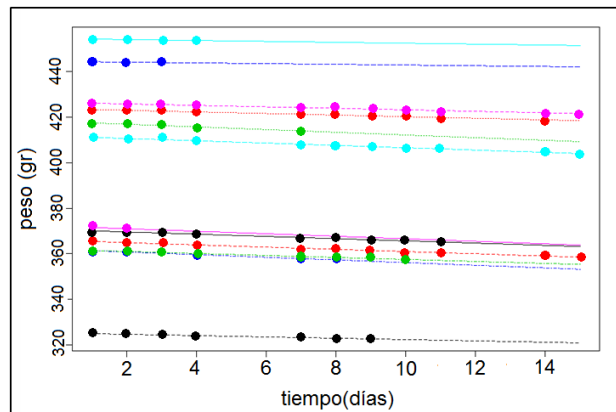


Gráfico 5.1: Funciones aproximadas de los pesos en gramos de doce papayas medidas en once ocasiones durante los quince días del experimento piloto.

Con el fin de determinar el número de réplicas que deberá tenerse en cada nivel del factor película, se realizó una experimentación piloto para tener resultados preliminares. Se estimaron las funciones de pérdida de peso por medio del método de imputación detallado anteriormente.

Con los datos resultantes y bajo los supuestos teóricos de pérdida de peso, por medio de simulación se procedió a determinar el número de replicas a utilizarse.

5.3. Simulación

Por medio de simulación matemática, con los resultados de la experimentación piloto se simularon 10000 experimentos en las cuales con un nivel de significancia α de 0.05 en cada una de las simulaciones, se realizaron la prueba F de FANOVA y la prueba T de diferencias de medias, esto probando con diferentes tamaños de réplicas (3, 4, 5, 10, 15 y 20) y variando los supuestos de los efectos entre niveles (0.50, 0.75, 1.00, 1.25 y 1.50 gramos). Al final se obtuvieron las Funciones de Poder respectivas para cada prueba.

Cada papaya fue generada como

$$Y_{rp}(t) = \hat{\mu}_p(t) + Z s_p(t),$$

donde $Y_{rp}(t)$ representa la función de pérdida de peso de la papaya r con tipo de película p , $\hat{\mu}_p(t)$ es la función promedio de pérdida de peso para papayas con tipo de película p , Z es una variable aleatoria Normal estándar y $s_p^2(t)$ el estimador insesgado de la función de la varianza poblacional. El algoritmo de simulación fue realizado y ejecutado en el software estadístico libre R.

5.3.1. Resultados de la simulación. En la tabla 2 se observan los resultados de la simulación. Se muestran los distintos valores que toma la Función de Poder para las pruebas F y T, para los diferentes tamaños de replicas y las diferentes distancias.

Con los resultados de las simulaciones se decidió utilizar tres replicas por cada nivel, ya que asumiendo una diferencia promedio de 1 gramo entre papayas con películas estándar y películas con aceites esenciales, se obtiene un aceptable poder de 0.6238 para la prueba T y un 0.9653 para la prueba F.

Aunque la potencia es baja para la prueba T la decisión se dio en base a las restricciones de espacio y uso de los equipos en los laboratorios. Se muestra a continuación la distribución de los frutos en las refrigeradoras. En el gráfico 5.2, se puede observar que existe una única posición por repisa donde se encuentran dos papayas; una de ellas va a ser cortada el primer día

del experimento, mientras que la restante va a ser objeto del análisis del color durante todo el experimento.

Tabla 2: Funciones de Poder de las pruebas F y T realizadas por medio de simulación, para diferentes tamaños de réplicas y distancias.

		d = 0.50	d = 0.75	d = 1.00	d = 1.25	d = 1.50
3	r F	0.8841	0.9309	0.9653	0.9839	0.9934
	t	0.2533	0.4323	0.6238	0.7907	0.9002
4	r F	0.9377	0.9675	0.9856	0.9958	0.9992
	t	0.3275	0.5484	0.7656	0.9055	0.9704
5	r F	0.9576	0.9827	0.9946	0.9987	0.9996
	t	0.3870	0.6470	0.8557	0.9586	0.9932
10	r F	0.9921	0.9987	0.9999	1.0000	1.0000
	t	0.6381	0.9075	0.9882	0.9998	1.0000
15	r F	0.9972	1.0000	1.0000	1.0000	1.0000
	t	0.7809	0.9793	0.9995	1.0000	1.0000
20	r F	0.9997	1.0000	1.0000	1.0000	1.0000
	t	0.8864	0.9972	1.0000	1.0000	1.0000

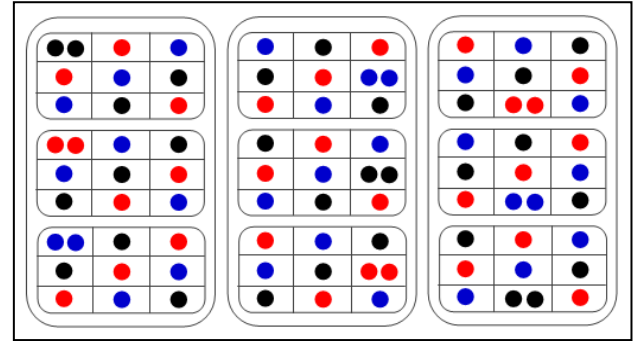


Gráfico 5.2: Posición espacial de las papayas en las refrigeradoras durante el experimento. Obsérvese las papayas sin película (negro), con película (rojo) y con películas con aceites esenciales (azul).

6. Análisis de Resultados

Dada la naturaleza de cada una de las variables analizadas, cada una de ellas fue modelizada por funciones con distintos sistemas bases. Véase la tabla 3.

Tabla 3: Sistemas bases seleccionados para modelizar las variables de interés.

Variable	Sistema Base	Parámetros
Peso	Exponencial	$\{ 1, e^{-0.1}, e^{-1}, e^{-3} \} \lambda = 10^{-4}$
Nivel pH	B-Spline	nbasis = 6, norder = 4
Sólidos Solubles	B-Spline	nbasis = 7, norder = 4
Dureza	B-Spline	nbasis = 7, norder = 4
Color	B-Spline	nbasis = 8, norder = 4, $\lambda = 0.479$

6.1. Resultados de Análisis de Varianza Funcional

Los resultados del Análisis Funcional de Varianza como de las pruebas funcionales de diferencias de medias se resumen para cada variable.

6.1.1. Peso. La función análoga al estadístico de prueba F tomó valores exorbitantes y como consecuencia el valor p de la prueba siempre fue cero durante la realización del experimento, es por esto que al rechazar la hipótesis nula concluimos que siempre existen diferencias significativas sobre el peso de la papaya, al aplicar algún tipo de película comestible, mas sin saber cuál de ellas dos es la más efectiva. El efecto de la aplicación la película con aceites esenciales resulta mayor únicamente durante el primer día del experimento; como resultado se rechazó H_0 para la prueba T de diferencias de medias durante este período de tiempo.

6.1.2. Nivel de pH. Alrededor del sexto día del experimento la Potencia de Explicación del Modelo alcanzó un 71.86% donde el valor p se encontró en zona de rechazo, dejando en evidencia que para este período de tiempo comprendido ente el quinto y séptimo día existen diferencias significativas en la acidez del fruto al aplicar películas comestibles de cualquier tipo.

El valor p correspondiente a la prueba unilateral de diferencia de funciones medias es siempre mayor a 0.10, por lo que en ningún momento del experimento se puede considerar favorable la aplicación de películas comestibles con aceites esenciales con el fin de preservar el nivel de pH en el fruto.

6.1.3. Sólidos Solubles. A partir del sexto día del experimento el valor p correspondiente a la prueba F de diferencias significativas es menor que 0.05 por lo que se rechaza H_0 que postula que no existen diferencias significativas en la aplicación de películas comestibles en las papayas, por lo que se concluye que sí existen diferencias significativas en la aplicación.

El análisis de efectos individuales es corroborado nuevamente al evidenciarse que jamás se rechaza H_0 que postula que el efecto de la aplicación de aceites esenciales en la película es nulo.

6.1.4. Dureza. Por medio de la función estadístico de prueba $F(t)$ y su respectiva función P_{VALUE} se confirma que durante todo el experimento hasta el día trece se encuentran diferencias significativas en la dureza del fruto al aplicársele películas comestibles.

La hipótesis nula de la prueba de hipótesis funcional que compara los dos tipos de película comestibles, jamás es rechazada por lo que se puede concluir que la aplicación de aceites esenciales no tiene efectos favorables en la dureza del fruto.

6.1.5. Color. Durante los primeros 6 días del experimento que tanto la función $F(t)$ y P_{VALUE} se encuentran en zona de rechazo por lo que únicamente en estos días (entre el noveno y décimo) es donde se encuentran diferencias significativas sobre el color, al aplicarse películas comestibles en los frutos.

Aplicando la prueba funcional de diferencias de medias encontramos que nunca se puede considerar que la aplicación de aceites esenciales en la película es favorable en el color del fruto, como indicador de la madurez del fruto.

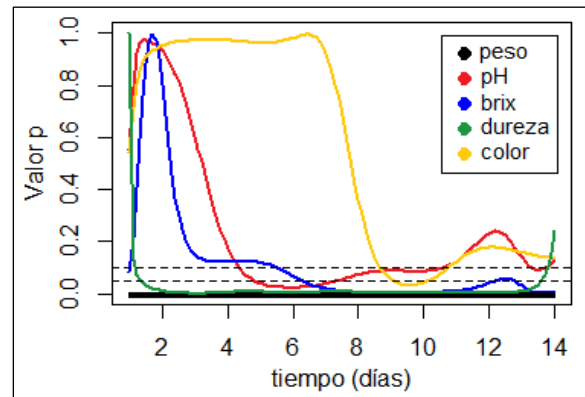


Gráfico 6.1: Función P_{VALUE} análoga al valor p o nivel de significancia de la prueba F de FANOVA con el fin de determinar diferencias significativas entre los niveles de los tratamientos.

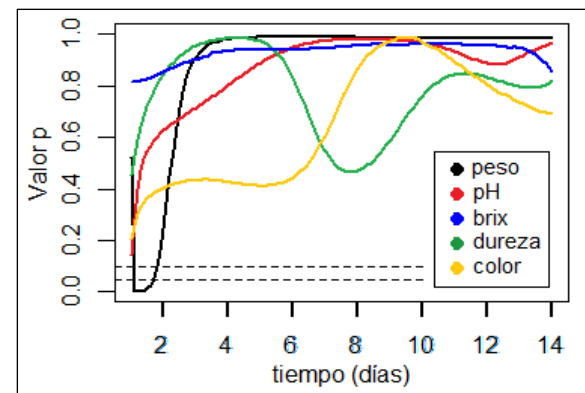


Gráfico 6.2: Función P_{VALUE} análoga al valor p o nivel de significancia de la prueba T funcional para diferencias de medias. Se comparan la película experimental con la estándar.

7. Conclusiones

- Se pudo comprobar la efectividad de las películas comestibles al retardar el proceso de maduración en las papayas, retrasando su pérdida de peso, previniendo el aumento del nivel pH, reduciendo los sólidos solubles, manteniendo la dureza y conservando la tonalidad oscura en el fruto.
- La aplicación de películas comestibles reduce significativamente la variabilidad en la pérdida de peso y en la dureza del fruto, por lo que si fuese de interés predecir la pérdida promedio de peso o dureza del fruto en un instante de tiempo t , de no aplicarse algún tipo de revestimiento se tendría mayor incertidumbre.
- La película comestible estándar resultó siempre mejor que la película comestible experimental (con aceites esenciales) al momento de reducir los cambios en los indicadores de interés que son evidencia de la maduración del fruto.
- No es de desacreditar la efectividad de las películas comestibles con aceites esenciales, pues si el lugar de almacenamiento del fruto es ciertamente húmedo, esto aumentará la probabilidad de que aparezcan pudriciones fungosas en las papayas y he aquí es donde la efectividad de esta película es válida, además que siempre será mejor en la conservación de la madurez del fruto comparado sino se aplicase ningún tipo de película.

8. Recomendaciones

- Disponer de los recursos necesarios (laboratorios, instrumentos, personal, equipos, etc.) para realizar el experimento anterior con un mayor número de réplicas para así obtener resultados con una mayor afijación.
- Dado que las mediciones no fueron realizadas a las mismas unidades experimentales durante el experimento porque los frutos eran cortados, se sugiere utilizar otros métodos para relacionar las unidades entre tiempo y tiempo; como sugerencia se podría considerar la parcela de la cual el fruto ha sido cosechado entre otros.
- Experimentar con otros agentes naturales en películas comestibles que eviten la contaminación fungosa del fruto y que sea igual o más efectiva en la conservación del fruto que una película comestible estándar.

9. Referencias

- [1] **Abramovich, F., Antoniadis, A., Sapatinas, T. y Vidakovic, B.** (2002). Optimal testing in functional analysis of variance models. Georgia Institute of Technology ISYE Statistics Technical Report.
- [2] **Almeida A., Reis J. D., Santos D., Vieira, T. y da Costa, M.** (2011), Study of preservation of papaya (Carica papaya L.) associated with the application of edible films, Universidade Federal de Sergipe (UFS), São Cristóvão, Sergipe (SE), Brasil.
- [3] **de Boor, C.** (2001), A Practical Guide to Splines, Springer series in statistics, New York - USA.
- [4] **Hoerl, Arthur E.** (1962), Application of ridge analysis to regression problems, Chemical Engineering Progress, 58, 54-59.
- [5] **Hooker, G.** (2010), Introduction to Functional Data Analysis, International Workshop on Statistical Modeling, Cornell University, Ithaca, NY - USA.
- [6] <http://www.functionaldata.org>, actualizado a Septiembre de 2009 y consultado a Septiembre de 2011.
- [7] <http://solidosolublesdenana.blogspot.com>, actualizado a Marzo de 2012 y consultado a Mayo de 2012.
- [8] http://www.statistics.com/index.php?page=glossary&term_id=766, actualizado a Agosto de 2011 y consultado a Septiembre de 2011.
- [9] **Kutner M., Nachtsheim C. y Neter J.**, Applied Linear Regression Models, 4th edition, McGraw-Hill Irwin, 2004.
- [10] **Press, W. H., Teukolsky, S. A., Vetterling, W. T. y Flannery, B. P.** (1999), Numerical recipes in C, Segunda edición, Cambridge, Cambridge University Press.
- [11] **Ramsay, J.O. y Silverman, B.W.** (2005), Functional data analysis, Segunda edición, Springer series in statistics, New York - USA.
- [12] **SAS/STAT(R)** (2012), Manual de Usuario, Versión 9.2, Segunda edición. Disponible en la web en http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_introcom_a00000000525.htm
- [13] **Satterwaite, F. E.** 1946. An approximate distribution of estimates of variance components. Biometrics Bull 2:110-4.
- [14] **Stein, Charles M.** (1956), "Inadmissibility of the usual estimator for the mean of a multivariate distribution", Proc. Third Berkeley Symp, 1:197-206.
- [15] **Tychonoff, Andrey N.** (1943), "Об устойчивости обратных задач [On the stability of inverse problems]". Doklady Akademii Nauk SSSR, 39 (5): 195-198.
- [16] **Zurita, G.** (2010). "Probabilidad y Estadística, Fundamentos y Aplicaciones", Segunda Edición, Instituto de Ciencias Matemáticas ESPOL, Guayaquil, Ecuador.