Departamento de Matemática Aplicada a las T.I.C.

ASIGNATURA: ESTADÍSTICA Y PROCESOS ESTOCÁSTICOS
EXAMEN FINAL (Otoño 2016-17) **Duración:** 3 horas
FECHA: 11 de Enero de 2017
Fecha publicación notas: 20 de Enero de 2017
Fecha revisión examen: 24 de Enero de 2017

SOLUCIONES

Ejercicio 1 *(1.5 puntos) Dos grupos de 9 personas cada uno deciden medir su altura. En el primer grupo resulta que sólo 4 personas miden menos de* 1.60 *metros (m) mientras que en el segundo grupo son* 2 *las personas que miden menos de* 1.60 *m.*

a) Si se eligen tres personas al azar del primer grupo, calcula la probabilidad de que dos o más midan menos de 1.60 *m.*

b) Supongamos que una de las personas del primer grupo se pasa al segundo grupo. Si se eligen dos personas de este segundo grupo modificado, calcula la probabilidad de que las dos midan menos de 1.60 *m.*

c) En las condiciones del apartado anterior y, sabiendo que las dos personas elegidas en el segundo grupo miden menos de 1.60 *m, calcula la probabilidad de que la persona que pasó del primer grupo al segundo mida* 1.60 *m. o más.*

Solución.

a) Definimos los sucesos

A: dos personas o más del primer grupo midan menos de 1.60 m
A_2 : dos personas del primer grupo midan menos de 1.60 m
A_3 : tres personas del primer grupo midan menos de 1.60 m

Entonces,

$$P(A) = P(A_2 \cup A_3) = \frac{\binom{4}{2} \cdot \binom{5}{1}}{\binom{9}{3}} + \frac{\binom{4}{3}}{\binom{9}{3}} = 0.4048$$

b) Definimos los sucesos

B_- : la persona que pasa del primer grupo al segundo grupo mida menos de 1.60 m
B_+ : la persona que pasa del primer grupo al segundo grupo mida 1.60 m o más
B : las dos personas que que se eligen del segundo grupo miden menos de 1.60 m

Para calcular $P(B)$ vamos a utilizar el teorema de la Probabilidad Total

$$P(B) = P(B_-) \cdot P(B/B_-) + P(B_+) \cdot P(B/B_-) = \frac{4}{9} \cdot \frac{\binom{3}{2}}{\binom{10}{2}} + \frac{5}{9} \cdot \frac{\binom{2}{2}}{\binom{10}{2}} = 0.04197$$

c) Usaremos ahora el teorema de Bayes para calcular

$$P(B_+/B) = \frac{P(B_+) \cdot P(B/B_+)}{P(B)} = \frac{\frac{5}{9} \cdot \frac{1}{45}}{0.04197} = 0.2941$$

Ejercicio 2 *(1.5 puntos) El diámetro de ciertos componentes, en cm., es una variable aleatoria X con función de densidad*

$$f(x) = \begin{cases} \dfrac{3}{4}x^2(2-x) & 0 \le x \le 2 \\ 0 & \textit{en el resto} \end{cases}$$

Suponiendo independencia entre los componentes

a) Halla la probabilidad de que el primer componente con diámetro superior a 1 cm. sea el quinto que se fabrica.

b) Si se fabrican 10 componentes, calcula la probabilidad de que más de 2 tengan diámetros inferiores o iguales a 1 cm.

c) Sabiendo que la varianza de X vale 0.16 cm^2, calcula aproximadamente la probabilidad de que el diámetro medio de 100 componentes, elegidos al azar, esté entre 1.1 cm. y 1.3 cm.

Solución.

a) La probabilidad de que un componente tenga un diámetro superior a 1 cm es

$$P(X>1) = \left[\frac{1}{2}x^3 - \frac{3}{16}x^4\right]_1^2 = \frac{11}{16} = 0.6875.$$

Se define la variable aleatoria

Y: número de componentes que se fabrican hasta que aparece el primero con diámetro superior a 1 cm.

$Y \sim Geo(0.6875)$

Se trata de calcular

$$P(Y=5) = 0.3125^4 \cdot 0.6875 = 0.00656$$

b) Se define la variable aleatoria

Z:número de componentes con diámetro menor o igual que 1 cm.

$Z \sim Bin(10, p = 0.3125)$

Se trata de calcular

$$P(Z>2) = 1 - P(Z \le 2) = 1 - P(Z=0) - P(Z=1) - P(Z=2) =$$

$$= 1 - 0.6875^{10} - \binom{10}{1} \cdot 0.3125 \cdot 0.6875^9 - \binom{10}{2} \cdot 0.3125^2 \cdot 0.6875^8 = 0.649855$$

c) Definimos las variables aleatorias

$$X_i : \text{“diámetro del } i\text{-ésimo componente”}, \qquad \text{con } i = 1..100,$$

Estas variables tienen la misma distribución que X, cuya media es μ y su desviación típica es $\sigma = \sqrt{0.16} = 0.4$. Se trata de calcular la probabilidad de que la variable

$$\overline{X} = \frac{X_1 + \cdots + X_{100}}{100}$$

esté entre 1.1 y 1.3. Se cumplen las condiciones para aplicar el Teorema Central del Límite. Por tanto,

$$\overline{X} = \frac{X_1 + \cdots + X_{100}}{100} \overset{aprox}{\sim} N\left(\mu, \sigma/\sqrt{100}\right)$$

Previamente tenemos que calcular el valor de μ. Entonces,

$$\mu = E[X] = \int_0^2 x \cdot f(x) = \int_0^2 \frac{3}{4} x^3 (2-x)\, dx = 1.2$$

Entonces,

$$P\left(1.1 \le \overline{X} \le 1.3\right) = P\left(\frac{1.1-1.2}{0.4/10} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{100}} \le \frac{1.3-1.2}{0.4/10}\right) \simeq P\left(-2.5 \le Z \le 2.5\right) =$$

$$= P\left(Z \le 2.5\right) - P\left(Z < -2.5\right) == 2P\left(Z \le 2.5\right) - 1 = 2 \cdot 0.9938 - 1 = 0.9876.$$

donde $Z \sim N(0,1)$.

Ejercicio 3 *(1 punto) Una universidad madrileña tiene un 50 % de estudiantes de la Comunidad de Madrid, un 40 % de estudiantes del resto de España y el resto son estudiantes extranjeros.*

a) Calcula la probabilidad de que, de 10 estudiantes seleccionados, 6 sean de Madrid y 2 sean extranjeros.

b) Calcula la probabilidad de que, de 15 estudiantes seleccionados, al menos 3 sean extranjeros.

***Solución*.**

a) Sean
X_1 : "número de estudiantes, entre 10 seleccionados, que son de Madrid".
X_2 : "número de estudiantes, entre 10 seleccionados, que son del resto de España".
X_3 : "número de estudiantes, entre 10 seleccionados, que son extranjeros".

Entonces, $(X_1,\ X_2,\ X_3) \sim Mult\ (n = 10,\ p_1 = 0.5,\ p_2 = 0.4,\ p_3 = 0.1)$

Si $X_1 = 6$ y $X_3 = 2$, necesariamente $X_2 = 2$, entonces

$$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6! \cdot 2! \cdot 2!} 0.5^6 \cdot 0.4^2 \cdot 0.1^2 = 0.0315$$

b) Sea Y la variable aleatoria "número de estudiantes, entre 15 seleccionados, que son extranjeros"
$Y \sim Bin(n = 15, p = 0.1)$

$$\begin{aligned} P(Y \ge 3) &= 1 - P(Y < 3) \\ &= 1 - \left[\binom{15}{0} 0.9^{15} + \binom{15}{1} 0.1^1 \cdot 0.9^{14} + \binom{15}{2} 0.1^2 \cdot 0.9^{13}\right] = 0.184 \end{aligned}$$

Ejercicio 4 *(1 punto) Sea* $(X,\ Y)$ *una variable aleatoria bidimensional con función de densidad conjunta*

$$f(x,y) = \begin{cases} \dfrac{1+x\cdot y}{4} & si\ |x|<1,\ |y|<1 \\ \\ 0\ en\ el\ resto \end{cases}$$

a) Estudia razonadamente si son X e Y independientes.

b) Calcula la probabilidad de que $X+Y$ sea superior a 1.

Solución.

a) La función de densidad conjunta es diferente de cero en el cuadrado $(-1,\ 1)\times(-1,\ 1)$. Para hallar las funciones de densidad marginales integramos la función de densidad conjunta en el cuadrado $(-1,\ 1)\times(-1,\ 1)$.
$f_X(x) = \displaystyle\int_{-1}^{1} \frac{1+x\cdot y}{4}\,dy = \frac{1}{4}\left(1+x\frac{1}{2}-(-1)-x\frac{1}{2}\right) = \frac{1}{2}$. Por tanto la función de densidad marginal de X es

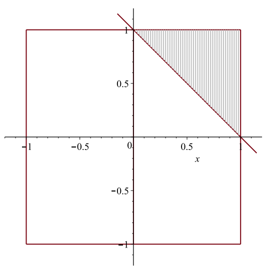
$$f_X(x) = \begin{cases} \dfrac{1}{2} & \text{si } -1<x<1 \\ \\ 0 \text{ en el resto} \end{cases}$$

por lo que X sigue una distribución uniforme en el intervalo $(-1,\ 1)$, es decir $X \sim U(-1,\ 1)$.
Razonando análogamente, $f_Y(y) = \displaystyle\int_{-1}^{1} \frac{1+x\cdot y}{4}\,dx = \frac{1}{4}\left(1+y\frac{1}{2}-(-1)-y\frac{1}{2}\right) = \frac{1}{2}$. Por tanto la función de densidad marginal de Y es

$$f_Y(y) = \begin{cases} \dfrac{1}{2} & \text{si } -1<y<1 \\ \\ 0 \text{ en el resto} \end{cases}$$

así que Y también sigue una distribución uniforme en el intervalo $(-1,\ 1)$, $Y\sim U(-1,\ 1)$.
Si X e Y fuesen independientes $f(x,y) = f_X(x)\cdot f_Y(y),\ \forall x,y\in\mathbb{R}$. En el cuadrado donde no es cero la función de densidad conjunta $\dfrac{1+x\cdot y}{4} \neq \dfrac{1}{2}\cdot\dfrac{1}{2}$, entonces X e Y no son independientes.

b) Para calcular la probabilidad de que $X+Y$ sea superior a 1 tenemos que integrar la función de densidad conjunta en el recinto sombreado de la siguiente figura



$$P(X+Y>1) = \int_0^1 dx \int_{1-x}^{1} \frac{1+x\cdot y}{4}\,dy = \frac{1}{4}\int_0^1 \left(y + x\frac{y^2}{2}\Big|_{1-x}^{1} \right) dx =$$

$$= \frac{1}{4}\int_0^1 \left(-\frac{x^3}{2}+x^2+x\right) dx = \frac{1}{4}\left(-\frac{x^4}{8}+\frac{x^3}{3}+\frac{x^2}{2}\Big|_0^1\right) = \frac{1}{4}\cdot\frac{17}{24} = \frac{17}{96} \backsimeq 0.1770833$$

Ejercicio 5 *(2 puntos) Se sabe que la anchura de las pistas de un circuito microstrip fabricado por una máquina sigue una distribución normal. Se ha obtenido una muestra de 16 pistas sabiendo que la anchura media y la cuasivarianza de la muestra son, respectivamente,* $\bar{x} = 192.1\mu m$ *y* $s_1^2 = 30\mu m$.

Nota 1 *Sea X una variable aleatoria con distribución t de Student con 15 grados de libertad. La siguiente tabla muestra los valores de su función de distribución en diferentes abscisas*

x	-2.13	-1.75	-1.34	1.34	1.53	1.75	2.13	2.6
$P(X \leq x)$	0.025	0.05	0.1	0.9	0.925	0.95	0.975	0.99

Nota 2 *Sea Y una variable aleatoria con distribución* χ^2 *con 15 grados de libertad. La siguiente tabla muestra los valores de su función de distribución en diferentes abscisas*

y	6.26	7.26	8.54	22.31	23.45	24.99	27.49	30.58
$P(Y \leq y)$	0.025	0.05	0.1	0.9	0.925	0.95	0.975	0.99

a) Calcula un intervalo de confianza al 80 % para la varianza de las anchuras de las pistas fabricadas por la máquina.

b) Contrasta la hipótesis de que la anchura media de las pistas es igual a 190 μm, considerando un nivel de significación $\alpha = 0.05$. *¿Para qué niveles de significación se rechazaría la hipótesis?*

Solución.

a)

$$P\left(q_{0.1} \leq \frac{(n-1)S_1^2}{\sigma^2} \leq q_{0.9}\right) = 0.8 \Longleftrightarrow P\left(\frac{(n-1)S_1^2}{q_{0.9}} \leq \sigma^2 \leq \frac{(n-1)S_1^2}{q_{0.1}}\right) = 0.8$$

siendo $q_{0.1}$ y $q_{0.9}$ los cuantiles 0.1 y 0.9 de una distribución χ^2_{15}

El intervalo de confianza para σ^2:

$$\left(\frac{15 \cdot 30}{22.31}, \frac{15 \cdot 30}{8.54}\right) = (20,17\ ,\ 52,69)$$

b) H_0: $\mu = 190$ H_1: $\mu \neq 190$

Si H_0 es cierta, $T = \dfrac{\bar{X} - 190}{S_1/\sqrt{16}} \sim t_{15}$

En la muestra $\bar{x} = 192,1 \quad s_1 = \sqrt{30}$

El valor de T en la muestra es

$$\frac{192,1 - 190}{\sqrt{30}/\sqrt{16}} = 1.53$$

Siendo q_α el cuantil de una t_{15}, el intervalo de aceptación es:

$$(q_{0,025}, q_{0,975}) = (-q_{0,975}, q_{0,975}) = (-2,13\ ,\ 2,13)$$

$1,53 \in (-2,13\ ,\ 2,13)$, luego se acepta H_0

También se puede hacer calculando el $p-valor$:

$$p - valor = 2 \cdot P(t_{15} > 1.53) = 2 \cdot 0.075 = 0.15 > 0.05 = \alpha$$

y, por lo tanto, se aceptará H_0. Finalmente, para todo $\alpha > p - valor = 0.15$ se rechazará la hipótesis.

Ejercicio 6 *(1.5 puntos) Sean $X(t)$ e $Y(t)$ dos procesos independientes, estacionarios en sentido amplio, con medias iguales a cero y la misma función de autocorrelación $R(\tau)$. Consideremos el proceso $Z(t) = 2X(t) - 5Y(t)$.*

a) Calcula la media y la función de autocorrelación de $Z(t)$. Determina si $Z(t)$ es estacionario en sentido amplio.

b) Si los procesos $X(t)$ e $Y(t)$ además son gaussianos y la función de autocorrelación común es $R(\tau) = e^{-|\tau|}$, halla las distribuciones de primer orden del proceso $Z(t)$.

c) Halla la distribución del vector aleatorio $\begin{pmatrix} X(0) \\ Y(0) \\ X(3) \\ Y(3) \end{pmatrix}$.

d) Con los supuestos de los apartados anteriores, halla la distribución de $\begin{pmatrix} Z(0) \\ Z(3) \end{pmatrix}$.

Solución.

a) Por ser el operador $E[\cdot]$ lineal,

$$\mu_Z(t) = E[Z(t)] = E[2X(t) - 5Y(t)] = 2\mu_X(t) - 5\mu_Y(t) = 0$$

Las funciones de autocorrelación de los procesos, $R_X(\tau)$ y $R_Y(\tau)$, coinciden con $R(\tau)$.

$$R_Z(t,\ t+\tau) = E[Z(t)Z(t+\tau)] = E[(2X(t) - 5Y(t)) \cdot (2X(t+\tau) - 5Y(t+\tau))] =$$

$$= E[4X(t)X(t+\tau) - 10X(t)Y(t+\tau) - 10Y(t)X(t+\tau) + 25Y(t)Y(t+\tau)]$$

Por ser los procesos $X(t)$ e $Y(t)$ independientes y estacionarios de media cero

$$R_Z(t,\ t+\tau) = 4R_X(\tau) - 10\mu_X\mu_Y - 10\mu_Y\mu_X + 25R_Y(\tau) = 4R_X(\tau) + 25R_Y(\tau) = 29R(\tau)$$

El proceso $Z(t)$ tiene la media, μ_Z, constante y su función de autocorrelación no depende de t, luego es estacionario en sentido amplio.

b) Para cada valor de t fijo, las variables aleatorias $X(t)$ e $Y(t)$ son independientes y gaussianas y, por tanto, su distribución conjunta es normal bidimensional con vector de medias $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y matriz de covarianzas $\begin{pmatrix} R_X(0) & 0 \\ 0 & R_Y(0) \end{pmatrix} = \begin{pmatrix} R(0) & 0 \\ 0 & R(0) \end{pmatrix}$.

Por tanto,

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

Como $Z(t) = \begin{pmatrix} 2 & -5 \end{pmatrix} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}$, se verifica que $Z(t)$ sigue una distribución normal.

Como ya hemos probado, $\mu_Z(t) = 0$.

$$Var[Z(t)] = \begin{pmatrix} 2 & -5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -5 \end{pmatrix} = 4 + 25 = 29$$

De otro modo, $Var[Z(t)] = R_Z(0) = 29R(0) = 29$

Así, $Z(t) \sim \mathcal{N}(0, \sigma_Z^2 = 29)$.

c) Como

$$\begin{pmatrix} X(0) \\ X(3) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & e^{-3} \\ e^{-3} & 1 \end{pmatrix}\right), \qquad \begin{pmatrix} Y(0) \\ Y(3) \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & e^{-3} \\ e^{-3} & 1 \end{pmatrix}\right)$$

y $[X(0), X(3)]$ es independiente de $[Y(0), Y(3)]$, se verifica que la distribución conjunta de las variables $X(0), Y(0), X(3), Y(3)$ es normal de dimensión 4, con vector de medias $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Además

$Var[X(t)] = Var[Y(t)] = R(0) = e^0 = 1$

$Cov[X(0), X(3)] = Cov[Y(0), Y(3)] = R(3) = e^{-3}$

$Cov[X(t_1), Y(t_2)] = 0$, por ser los procesos $X(t)$ e $Y(t)$ independientes.

Con lo que la matriz de covarianzas de $[X(0), Y(0), X(3), Y(3)]$ es

$$\begin{pmatrix} R_X(0) & 0 & R_X(3) & 0 \\ 0 & R_Y(0) & 0 & R_Y(3) \\ R_X(3) & 0 & R_X(0) & 0 \\ 0 & R_Y(3) & 0 & R_Y(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 & e^{-3} & 0 \\ 0 & 1 & 0 & e^{-3} \\ e^{-3} & 0 & 1 & 0 \\ 0 & e^{-3} & 0 & 1 \end{pmatrix}$$

De manera que,

$$\begin{pmatrix} X(0) \\ Y(0) \\ X(3) \\ Y(3) \end{pmatrix} \sim \mathcal{N}_4 \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & e^{-3} & 0 \\ 0 & 1 & 0 & e^{-3} \\ e^{-3} & 0 & 1 & 0 \\ 0 & e^{-3} & 0 & 1 \end{pmatrix} \right).$$

d)

$$\begin{pmatrix} Z(0) \\ Z(3) \end{pmatrix} = \begin{pmatrix} 2 & -5 & 0 & 0 \\ 0 & 0 & 2 & -5 \end{pmatrix} \begin{pmatrix} X(0) \\ Y(0) \\ X(3) \\ Y(3) \end{pmatrix}$$

$$\begin{pmatrix} E[Z(0) \\ E[Z(3)] \end{pmatrix} = \begin{pmatrix} 2 & -5 & 0 & 0 \\ 0 & 0 & 2 & -5 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

La matriz de covarianzas de $[Z(0), Z(3)]$ es

$$\begin{pmatrix} 2 & -5 & 0 & 0 \\ 0 & 0 & 2 & -5 \end{pmatrix} \begin{pmatrix} 1 & 0 & e^{-3} & 0 \\ 0 & 1 & 0 & e^{-3} \\ e^{-3} & 0 & 1 & 0 \\ 0 & e^{-3} & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -5 & 0 \\ 0 & 2 \\ 0 & -5 \end{pmatrix} = \begin{pmatrix} 4 + 25 & 4e^{-3} + 25e^{-3} \\ 4e^{-3} + 25e^{-3} & 4 + 25 \end{pmatrix}$$

Con lo que,

$$\begin{pmatrix} Z(0) \\ Z(3) \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 29 \cdot \begin{pmatrix} 1 & e^{-3} \\ e^{-3} & 1 \end{pmatrix} \right)$$

Ejercicio 7 *(1.5 puntos) El número de peticiones que recibe un servidor en cada intervalo de tiempo $[0, t)$ es un proceso de Poisson $N(t)$ de tasa 6 peticiones por segundo.*

a) Calcula la probabilidad de que en un segundo no haya ninguna petición.

b) Calcula la probabilidad de que en el primer segundo haya más de 1 petición y que en el siguiente segundo haya más de 2 peticiones.

c) *Si en el transcurso de un determinado segundo ya ha habido 2 peticiones ¿cuál es la probabilidad de que al finalizar ese segundo haya menos de 4 peticiones?*

Solución.

a) Para cada valor de t fijo, $N(t) \sim Pois(6t)$, $\qquad P[N(t) = i] = e^{-6t} \frac{(6t)^i}{i!}, \qquad i = 0, 1, 2, \ldots$.

El número de peticiones que se producen en un segundo se puede expresar como $N(t+1) - N(t)$, $t > 0$. $N(t+1) - N(t)$ tiene igual distribución que $N(1)$. es decir $N(1) \sim Pois(6)$. Finalmente, la probabilidad pedida es: $\qquad P(N(1) = 0) = e^{-6} \simeq 0.0025$.

b) La probabilidad pedida viene dada por $P(N(t+1) - N(t) > 1,\ N(t+2) - N(t+1) > 2)$. Tenemos que $[N(t+1) - N(t)]$ y $[N(t+2) - N(t+1)]$ son variables aleatorias independientes con la misma distribución que $N(1)$. Por otro lado, tenemos que:

$$P(N(1) > 1) = 1 - [P(N(1) = 0) + P(N(1) = 1)] = 1 - e^{-6}(1 + 6) \simeq 0.9826$$

$$P(N(1) > 2) = 1 - [P(N(1) = 0) + P(N(1) = 1) + P(N(1) = 2)] = 1 - e^{-6}\left(1 + 6 + \frac{6^2}{2!}\right) \simeq 0.938$$

Se verifica que

$$\begin{aligned} P(N(t+1) - N(t) > 1, N(t+2) - N(t+1) > 2) &= \\ = P(N(t+1) - N(t) > 1) \cdot P(N(t+2) - N(t+1) > 2) &= \\ = 0.9826 \cdot 0.938 = 0,9217 & \end{aligned}$$

c)

$$P(N(1) < 4/N(1) \geq 2) = \frac{P(2 \leq N(1) < 4)}{P(N(1) \geq 2)} = \frac{P(N(1) = 2) + P(N(1) = 3)}{P(N(1) > 1)} \simeq \frac{0.1339}{0.9826} = 0.1362$$