SOLUCION TAREA MA34B

Profesor: Alexis Peña. Auxiliar:Diego Díaz Espinoza

June 3, 2007

1. Los siguientes datos corresponden a la concentración disuelta (mg/L) de carbón orgánico recolectado de abono orgánico:

22.74	29.80	27.10	16.51	6.51	8.81	5.29	20.46	14.90	33.67
30.91	14.86	15.91	15.35	9.72	19.80	14.86	8.09	17.90	18.30
5.20	11.90	14.00	7.40	17.50	10.30	11.40	5.30	15.72	20.46
16.87	15.42	22.46							

a) Construya un intervalo de confianza al 99% para la media de la concentración de carbón orgánico disuelto. Interprete el intervalo.

Asumiendo que el experimento de tomar la concentración de carbón orgánico recolectado de abono orgánico tiene un error que se distribuye $N(0, \sigma^2)$ se considera que cada dato de la muestra se distribuye $N(\bar{x}, \frac{\sigma^2}{n})$. Luego el intervalo de confianza viene dado por: $\bar{x} \pm t_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ donde $\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$. Luego de este análisis se puede comprobar usando un software de estadística, en este caso R (http://www.r-project.org). el que arroja el siguiente resultado con una confianza del 99% $[12.4, 19.4]$. A continuación el output de R:

```
> #PREGUNTA 1, PARTE A
> A<-scan(file="pregunta_1_a_tarea.txt",what=0);
Read 33 items
> #Varianza
> var(A);
[1] 54.50644
> #Desviacion Estandar
> sd(A);
[1] 7.382848
> #Promedio
> mean(A);
[1] 15.92182
```

```
> #P(N>n)
> qnorm(0.005);
[1] -2.575829
> #Test T
> t.test(F, conf.level=0.99);

        One Sample t-test

data:  F
t = 12.3887, df = 32, p-value = 9.422e-14
alternative hypothesis: true mean is not equal to 0
99 percent confidence interval:
 12.40235 19.44128
sample estimates:
mean of x
 15.92182

>
```

2. Un segundo conjunto de datos representa la concentración de carbón orgánico (mg/L) recolectado de abono mineral.

8.50	3.91	9.29	21.00	10.89	10.30	11.56	7.00	3.99	3.79
5.50	4.71	7.66	11.72	11.80	8.05	10.72	21.82	22.62	10.74
3.02	7.45	11.33	7.11	9.60	12.57	12.89	9.81	17.99	21.40
8.37	7.92	17.90	7.31	16.92	4.60	8.50	4.80	4.90	9.10
7.90	11.72	4.85	11.97	7.85	9.11	8.79			

b)Construya un intervalo al 99% de confianza para la media de la concentración disuelta de carbón orgánico recolectado de abono mineral. Interprete el intervalo.

La solución es la misma que a pero cambian los datos, luego el intervalo de confianza viene dado por:

$[8.1, 11.9]$

El output de R arroja:

```
> #PREGUNTA 1, PARTE B
> B<-scan(file="pregunta_1_b_tarea.txt",what=0);
Read 47 items
> #Varianza
> var(B);
[1] 24.78904
> #Desviacion Estandar
> sd(B);
[1] 4.978859
```

```
> #Promedio
> mean(B);
[1] 10.02660
> #P(N>n)
> qnorm(0.005);
[1] -2.575829
> #Test T
> t.test(B, conf.level=0.99);

        One Sample t-test

data:  B
t = 13.8062, df = 46, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
99 percent confidence interval:
  8.075176 11.978016
sample estimates:
mean of x
 10.02660

>
```

c) Compare al 99% de confianza los intervalos para las concentraciones de carbón orgánico en abono mineral y orgánico. Existen evidencias suficientes para afirmar que ambas concentraciones son similares?.

Para la comparación de dos medias se usa el test $H0 : \mu_1 = \mu_2$o equivalentemente $H0 : \mu_1 - \mu_2 = 0$ Con esto y asumiendo normalidad en ambas muestras se encuentra que la región de rechazo del test viene dada por:

$[t/t > T, T = T_{(\frac{\alpha}{2}, n1+n2-2)}]$

Con $n1 = 23$, $n2 = 47$, $\frac{\alpha}{2} = 0.005$
Luego la región de rechazo viene dada por:

$[t/t > 2.66]$

Y para calcular t se usa:

$\frac{\bar{y1} - \bar{y2}}{s\sqrt{\frac{1}{n1} + \frac{1}{n2}}}$

Con

$s = \sqrt{\frac{(n1-1)s_1^2 + (n2-1)s_2^2}{n1+n2-2}} = 5.87$

Donde S_1y S_2 fueron calculados en a y b respectivamente.

Con esto se obtiene que $t = 4.04$Con lo que se rechaza la hipótesis nula. Luego no hay evidencia suficiente para afirmar que ambas concentraciones son similares.

Comprobando con R, se obtiene el siguiente output:

```
> #PREGUNTA 1, PARTE C
> A<-scan(file="pregunta_1_a_tarea.txt",what=0);
Read 33 items
> B<-scan(file="pregunta_1_b_tarea.txt",what=0);
Read 47 items
> t.test(A,B, mu=0,var.equal=TRUE, conf.level=0.99);

        Two Sample t-test

data:  A and B
t = 4.2685, df = 78, p-value = 5.488e-05
alternative hypothesis: true difference in means is not equal to 0
99 percent confidence interval:
 2.248631 9.541813
sample estimates:
mean of x mean of y
 15.92182  10.02660
```

En donde además se puede comprobar que en el intervalo de confianza no está el cero, luego no se puede aceptar la hipótesis nula.

3. Un laboratorio farmaceútico adquiere un determinado componente de dos diferentes proveedores. El nivel medio de impureza en el componente en bruto es aproximadamente el mismo para ambos proveedores, pero la cadena está preocupada por la variabilidad de dicha impureza debido a los transportes de los mismos. Si el nivel de impureza tiende a variar excesivamente desde un lugar a otro esto puede afectar la calidad del producto farmaceútico. Para comparar dicha variabilidad - en porcentaje - de la impuerza para los dos proveedores, el laboratorio selecciona 10 envíos de cada uno de los proveedores y les mide dicho porcentaje. Las medias de las muestras y sus varianzas se muestran acontinuación:

Proveedor A	Proveedor B
$\overline{x_1}$=1.89	$\overline{x_2}$=1.89
$S_1^2 = 0.273$	$S_1^2 = 0.094$
$n_1 = 10$	$n_2 = 10$

a) Son los datos obtenidos suficientes para indicar que existe una variabilidad en la impureza de los componentes debido a los traslados para los dos proveedores?. Use $\alpha = 0.10$. Basado en los resultados del test, qué recomendaría ud. al laboratorio?

La hipótesis nula queda como:

$H_0 : \sigma_1 = \sigma_2$

Se usa la F de Fisher como test. Luego la región crítica viene dada por:

$[f/f > F, F = F_{(\frac{\alpha}{2}, n1-1, n2-1)}]$

Con

$F = 3.18$

Luego la región de rechazo queda:

$[f/f > 3.18]$

Al calcular el valor obtenido del experimento se tiene:

$f = \frac{S_1^2}{S_2^2} = 2.9$

Luego no se rechaza la hipótesis nula. Notar que se está diciendo que existe evidencia para afirmar que las varianzas son iguales aún cuando existe mucha variabilidad en las muestras del laboratorio uno.