

dissabte, 7 gener 2012, 17:47

Lloc: **AulaVirtual - Aula Virtual de la Universitat Jaume I**  
 Curs: **IG23 - Ampliació d'Estadística (A) (2010/2011) (IG23-2010/2011)**  
 Glossari: **Dubtes preguntats altres anys**

## TEMA 1

### Qüestió: **Apareadas**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

#### **Resposta:**

En el cas de mostres emparellades no sé com calcules la  $s_d$ , per exemple en l'exercici 4 del examen del 14/9/2006.

Simplement has de fer la resta de les dades Algoritmo vigente - nuevo:  $8-0.7$ ,  $24.7-0.8$ , etc, i amb eixes dades traus la mitjana i la desviació típica ( $s_d$ ) amb la calculadora.

En la pàgina 50 del llibre tens també altre exemple resolt.

En l'examen 25/1/06, exercici 4 d'eixe mateix any, un altre. Primer has de traure les diferències entre Algoritmo 1 i 2, que dona:  $5.6-5.4=0.2$ ,  $3.8-3.8=0$ ,  $-0.8, 0.2, -0.5$ ,  $-0.1$ . Amb eixes dades (COMPTE :!!!No canvis signes!!!) traus la mitja ( $-0,166667$ ) i desviació típica ( $s_d$ ).

### Qüestió: **Calculadora**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

#### **Resposta:**

Com calcularia  $s$ , la desviació típica mostral, d'una sèrie de dades, amb la calculadora (Casio fx-85MS)? Jo inicie el mode estadístic i vaig teclejant dada,  $M+$ , dada  $M+$ . Per a calcular la mitja, tecleje shift + 2 + 1, però no sé traure d'ací la desviació típica.

En la pàgina 9 del tema 0 tens la fórmula (també al formulari). Amb la calculadora, una vegada introduïdes les dades (compte de no tindre introduïdes dades d'altre problema anterior, per no mesclar dades), tindràs una tecla amb  $\sigma_{N-1}$ , com fica a la pàgina 9 del tema 0.

En la teua calculadora, hauràs de tecletajar shift + 2 + 3 (al 3 pense que fica la  $\sigma_{N-1}$ , que és la nostra  $s$ ). Comprova-ho en un problema on sàpigues que dona de resultat, per assegurar-te.

### Qüestió: **F Snedecor**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

#### **Resposta:**

¿Cómo se saca  $F_{1-\alpha/2}$ ?

Por ejemplo, en el EJERCICIO 2 DEL EXAMEN 21/01/09, como nos dicen 95% confianza, entonces  $\alpha = 0.05$ . Por tanto,  $F_{1-\alpha/2} = F_{1-0.05/2} = F_{1-0.025} = F_{0.975}$  con  $(N_1 - 1, N_2 - 1) = (6-1, 5-1) = (5, 4)$  grados de libertad (g.l.). Si te fijas no hay tabla para  $F_{0.975}$ , así que hay que usar la propiedad que aparece en la primera tabla de la F de Snedecor (la de probabilidad 0.05, la que está al lado de la de la t-Student). Por ejemplo, si hay que obtener  $F_{0.975}$  con  $(5, 4)$  grados de libertad, es lo mismo que  $1/F_{0.025}$   $(4, 5)$  grados de libertad, fíjate que: 1) invertimos el valor, 2) la probabilidad es la complementaria, 3) el orden de los grados de libertad se intercambian.

Por eso es  $F_{0.975}(5, 4) = 1/7.39$ . Tienes que mirar la misma tabla que usaste para sacar el valor del otro extremo del intervalo, pero intercambiar el orden de los grados de libertad

y luego invertir ese valor.

En las soluciones de la autoevaluación del tema 1 puedes ver resuelto por completo un problema así: el problema 2 a).

#### Qüestió: **Interval de confiança**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

##### **Resposta:**

L'interval es compara amb el 0?

En els intervals de confiança quan hem de comparar 2 mitjanes, com és diferència de mitjanes, mirem si el 0 està o no dins de l'interval. Si està dins, les donem com mitjanes iguals, si no pertany el 0 a l'interval, les mitjanes poblacionals seran distintes. Fixat't que si són iguals, la seua resta (diferència) és zero, per això mirem el 0.

També mirem el 0 de la mateixa manera en el interval de diferència de proporcions.

En canvi, quan estem gastant l'interval de confiança del quocient de variàncies, es mira l'1, perquè si són iguals, el seu quocient (divisió) donarà 1. Amb aquest raonament, si l'1 està dins de l'interval, considerarem les variàncies iguals, sinò si 1 està fora de l'interval, les variàncies poblacionals seran distintes.

En la resta d'intervals que hem vist com són per a un paràmetre, no hi ha comparació i ja no es mira cap número.

En el llibre que us deixí a l'aulavirtual, tens exemples completament resolts, pas a pas, sobre interpretació d'intervals, d'exàmens passats.

#### Qüestió: **Media muestral y poblacional**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

##### **Resposta:**

¿Qué diferencia hay entre x barra y mu?

Una cosa es la media muestral que podemos obtener con los datos, por ejemplo medimos la altura de una muestra de 100 españoles y sacamos su media, esa es la media muestral (x barra), que podemos obtener con la calculadora.

Nosotros queremos poder extraer conclusiones, generalizar eso a toda la población, no quedarnos sólo con esas observaciones, queremos ver qué pasa con la mu. La mu es la media poblacional, la que tendríamos si midiéramos a TODOS los españoles, lo cual es bastante difícil, por no decir, imposible.

No conoceremos mu nunca, pero en base a la muestra podremos concluir algo sobre ella, bien con un intervalo de confianza, bien con un contraste de hipótesis.

#### Qüestió: **N (proporción)**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

##### **Resposta:**

En el problema 3 del examen 1/9/2008 apartado b), cuando vi la fórmula  $N = (...) \leq 1/4 (...)$  de la página 38 (abajo), interpreté el " $\leq$ " como que había que tomar el valor entero mas próximo a la baja. Es decir, si  $N \leq 756.25$ , como suponía que ese era el mayor valor posible, cogí el inferior entero, así cumplía que fuera entero y menor que el máximo posible.

En primer lugar, en el enunciado dice que uses la información de los datos anteriores, por tanto, el resultado correcto es 592.

Si no usaras la información que te dan (COSA que no dice el enunciado y por tanto, no se pide, y no debe confundirse con lo que se pide), entonces el redondeo sería a 757, pues 756.25 sería el mínimo número de datos a tomar. Pero, repito que eso no era lo que se

pedía. La respuesta correcta era 592.

En la página 52 del libro tienes un problema completamente resuelto sobre esto.

**Qüestió: t-Student**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

¿Cómo se sacan los grados de libertad para la tabla t-Student? Caso concreto del apartado b del 2º ejercicio del tema 1 de la Autoevaluación.

Los grados de libertad vienen en la propia fórmula del caso que estés usando, en la fórmula del caso F en este apartado. Si no están en la tabla los grados de libertad exactos que te salen, coge los más cercanos. En la tabla de la t de Student en la primera fila están las probabilidades y en la primera columna, los grados de libertad.

En la parte de arriba de cada tabla aparece como leer la tabla.

**Qüestió: Varianza poblacional y muestral**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** Duda examen 26/1/2005 ej3, c) No entiendo por qué en la solución a éste ejercicio eliges el caso B), cuando ya conocemos la varianza del apartado a).

Respuesta: ¡¡¡NO conoces la varianza POBLACIONAL ( $\sigma^2$ )!!!, sino que tienes la varianza muestral ( $s^2$ ). Fíjate que además el caso A) de los intervalos es de información complementaria. Una cosa es la varianza muestral que podemos obtener con los datos, por ejemplo medimos la altura de una muestra de 100 españoles y sacamos su varianza, esa es la varianza muestral ( $s^2$ ), que podemos obtener con la calculadora o el ordenador. La  $\sigma^2$  es la varianza poblacional, la que tendríamos si midiéramos a TODOS los españoles y sacáramos la varianza, lo cual es bastante difícil, por no decir, imposible. Con los intervalos y contrastes sí podremos decir algo acerca de ella.

En el tema 0, punto 0.3. tienes la diferencia entre muestra y población que era un punto de la IG12.

**Qüestió: ¿Qué caso usar?**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En intervalos de confianza, si en general las condiciones de un CASO no contempla algunas de las condiciones del enunciado del problema, ¿lo puedo usar igualmente?

Si cumple las condiciones del caso, puedes usarlo, aunque en el enunciado haya más información. Sin embargo, si alguna de las condiciones de un caso, no se contemplan en el enunciado, entonces, no puedes usar ese caso, y tendrías que usar otro caso. Siempre, os pondré una pregunta en la que podáis usar un caso que hayamos visto.

## TEMA 2

**Qüestió: Cas D i E i graus de llibertat**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En la part de autoevaluacio/problemas resueltos els exercicis que esta fotocopiats aparentment de un llibre usen unes formules que a vegades no apareixen al formulari, és més, per ser més concret en el 9-5 del tema 2 en rproblemes resoltts, diu " suponemos varianzas iguales" i si no me equivoco gasta la formula E) "varianzas desconocidas y desiguales" i en el 9-6 torna a ocurrer el mateix gasta la E) quan segons el enunciat tindria que usar la D) , així que m'agradaria saber en que me equivoco i per que o si per contra es un error del llibre o del formulari. Gràcies.

A més, en el 9-6 el que no entenc és perquè  $n_1$  a l'hora de fer els graus de llibertat posa  $n_1=15$   $n_2=10$  quan en les dades ens posa es  $n_1=5$   $n_2=10$ .

Està bé el llibre i el formulari, sols és que les fórmules estan escrites d'altra manera equivalent, però si les desenvolupes algebraicament, voràs que són les mateixes, donen el mateix resultat.

En el 9.5 és en variances iguals (gasta D del formulari i en el les fotocòpies d'eixe llibre també, encara que escrit d'altra forma), i en el 9.6 en variances diferents (gasta E del formulari i en el llibre fotocopiats també, encara que escrit d'altra forma), compte que en 9.6 fica que NO se desea varianzas iguales!

Respecte a l'últim dubte, sembla que en l'enunciat s'han enganyat, i en realitat  $n_1=15$ .

#### Qüestió: **Contrast d'hipòtesi**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** Com van els contrastos?

1) Primer es planteja el contrast, la  $H_0$  i  $H_1$ . En altra entrada explique com. 2) S'identifiquen les dades. 3) S'escollix el cas a gastar i 4) es substituïxen les dades en l'estadístic de contrast. 5) Es calcula la regió crítica en base a l'hipòtesi alternativa ( $H_1$ ) que es tinga (en altra entrada està explicat). 6) Per últim, es mira si el valor obtés al pas 4, pertany a la regió crítica o no. Si pertany a la regió crítica, rebutgem (rechazo)  $H_0$  i em quede amb  $H_1$ , si no pertany a la regió crítica, no rebutgem  $H_0$  i em quede amb  $H_0$ .

En el llibre tens també pas a pas un fum d'exemples d'exàmens d'altres anys.

#### Qüestió: **Contraste para tablas de contingencia**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En el problema 1 apartado b del examen de enero de 2009, cuando ya tengo el resultado,  $X^2=34.62$ , ¿cómo interpretas que rechazas  $H_0$ ? ¿Por qué 34.62 pertenece a la región crítica? ¿Cómo haces para definir  $H_0$ ? ¿Cómo eliges r y c?

Sí, como 34.62 pertenece a la región crítica (5.99, infinito) entonces rechazas  $H_0$ .

Es un contraste para tablas de contingencia y nos preguntan en el enunciado que si son homogéneas las distribuciones, por tanto,  $H_0$ : las distribuciones son homogéneas, y  $H_1$ : no son homogéneas. Los contrastes que hemos visto para  $X^2$  (chi cuadrado, punto 2.3 de los apuntes), son "fijos", en el sentido que si nos preguntan por homogeneidad, siempre irá en  $H_0$ , o si es por independencia en  $H_0$ . Mira el último párrafo del punto 2.3 de los apuntes.

Respecto a tu última pregunta, r y c son el número de filas y columnas de la tabla de contingencia.

Mira el libro que os dejé en el aulavirtual (páginas 72 a 75), donde tienes más ejemplos resueltos paso a paso, además de la autoevaluación del tema 2 (ejercicio 3).

#### Qüestió: **Estadístico de contraste**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En el apartado ii de a) del problema 2 del examen de 21/1/2009, pones que se utiliza el contraste D y que  $t=1.07$ , ¿podrías desarrollarme ese resultado?, no lo veo.

Aplicamos la fórmula del contraste D y sustituimos los valores (quizá te hayas equivocado calculando):  $media_1= 1,97833$ ,  $media_2= 1,754$ , valor que contrastamos  $=0$ ,  $s_1=0.264$ ,  $s_2=0.426$ ,  $N_1=6$ ,  $N_2=5$

$$1,97833 - 1,754 - 0$$

$$t = \frac{\text{raíz}((5 \cdot 0.264^2 + 4 \cdot .426^2)/(6+5-2))}{\text{raíz}(6 \cdot 5 / (6+5))}$$

**Qüestió: Hipòtesi alternativa**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** H\_0: diferencia de mitjes igual a 0

H\_1: diferencia de mitjes distint de 0

i

H\_0:  $\mu_1 \geq \mu_2$ H\_1:  $\mu_1 < \mu_2$ 

Significa el mateix?

**NO**, no és el mateix. En el primer cas, simplement estàs mirant si hi ha diferència de mitjanes (distint de zero). En el segon cas, està mirant-se si la mitjana de la primera població és menor que la de la segona població. Fixa't que m'he fixat en l'hipotesi alternativa (H\_1) per saber que estava comprovant.

Tot dependrà de que pregunten a l'enunciat, si simplement diferixen o si volen determinar si una és major o menor que l'altra.

De nou, en el llibre que us deixí a l'aualvirtual tens problemes d'altres anys resoltos amb les explicacions, per si vols fer-los una ullada

**Qüestió: P-valor**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En l'autoevaluació del tema 1, exercici 3, quan es fa la valoració mirant p-valor i quan mirant el interval de confiança?

Són equivalents en aquest cas (lilig l'observació de la pàgina 90 del llibre). Si no t'indique cap cosa, com preferixes. Si t'indique mitjançant una o altra, doncs això.

**Qüestió: Poisson**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** En l'exercici 4 de l'examen del 6/4/2004 a les solucions no ix com has calculat la freqüència esperada. Apareix la graella amb  $P(X=0) \cdot 576 = 0.395 \cdot 576 = 227,5358$ . He estat una estona pensant com s'obté però no he trobat solució.

Com s'està mirant si és Poisson, es gasta la funció de probabilitat de la Poisson (està al formulari), que vegeu en IG12:  $P(X=x) = e^{(-\lambda)} \cdot \lambda^x / x!$

La lambda està estimada abans (amb la mitjana, com es veugué al tema 1 en estimació puntual), sols has d'anar substituint cada valor de x.

En les pàgines 70-72 del llibre tens un idèntic resolt.

També en els problemes resoltos del tema 2, el problema 5 de l'autovaluació.

**Qüestió: Qui és H\_1?**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

Tinc dubtes a l'hora de desenvolupar les hipòtesi, ja que moltes vegades no sé perquè és eixa hipòtesi o una altra. Per exemple, en el apartat 2b) del examen de setembre de 2004.

Sempre has de ficar en H\_1 el que vols demostrar. Com vols mirar si el tiempo de respuesta medio con FIFO es mayor que con LIFO, has de ficar això a H\_1. O siga, en H\_1 anirà  $\mu_{FIFO} > \mu_{LIFO}$ , que és el mateix que  $\mu_{FIFO} - \mu_{LIFO} > 0$ . En la H\_0 sempre tens igualtat (quan estigues amb mitjanes, variances i proporcions), i has de ficar el mateix número que fiques en H\_1, en aquest cas igual a 0.

De nou, en el llibre que us deixí a l'aulavirtual, tens problemes d'altres anys resoltos amb les explicacions, per si vols fer-los una ullada.

Per altra part, en els contrastos no paramètrics, com tots el de la  $\chi^2$ , eixos tenen fixada la  $H_0$  i  $H_1$ , ahí no tindràs que calfar-te el cap. Per exemple, sempre serà, si vols contrastar independència, per exemple:

$H_0$ : les 2 variables són independents

$H_1$ : les 2 variables no són independents.

#### Qüestió: **Región crítica**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

##### **Resposta:**

En el ejercicio 2.4 del tema 2 de los apuntes, Contrastes de hipótesis, aplico el metodo D de contraste y en el momento de elegir la region crítica me surge la duda. La T me da -0.9739 y la t-student puede ser  $t(0.05,20)=1.725$  o  $t(0.025,20)=2.086$  según la region que escoja. Resulta que los intervalos que se me quedan son... (-infinito, -1.725) y (-infinito, -2.086) union (2.086, infinito). No sé cual de los dos se debería poner.

**¡CUIDADO!** La región crítica se escoge según la hipótesis alternativa  $H_1$  que tengas, que será el primer paso del contraste. En los contrastes paramétricos (punto 2.2), podrás elegir entre 3 posibles  $H_1$ . En la  $H_1$  pondrás aquello que quieras demostrar, que es lo que te indicará el enunciado. En este ejercicio, te pregunta si hay diferencia, por tanto, la  $H_1$  tiene que ser la de distinto. Con ello, la región crítica (sólo una, es una una ÚNICA) que tienes que escoger es la que corresponde a la  $H_1$  de distinto, que es la de la unión de los 2 intervalos: (-infinito, -2.086) U (2.086, infinito).

En el libro que os dejé en el aulavirtual hay muchos ejemplos de problemas de exámenes de otros años resueltos paso a paso, por si quieres echarles un vistazo.

### TEMA 3

#### Qüestió: **ARL**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

##### **Resposta:**

El apartado e del ejercicio 3 del examen 21/1/09, no lo veo claro.

El apartado e dice:

¿Cuál sería el número medio de muestras a representar hasta detectar el desplazamiento?

Esto se obtiene con la  $ARL = 1/p = 1/0.0057 = 175.4$

En la autoevaluación del tema 3 (ejercicio 2) tienes otro problema completamente resuelto sobre esto, por si quieres ver más. También en el libro que os dejé en el aulavirtual hay de exámenes de otros años.

#### Qüestió: **Desviació típica en un procés**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

##### **Resposta:**

Com estime la desviació típica d'un procés?

Una vegada està controlat el procés, és el rang mig/d<sub>2</sub>. El valor d<sub>2</sub> està a la taula de factors de control de qualitat, segons el nombre de mostretes ( n ) que s'agafen cada volta. Per exemple, en el problema 2 de l'autoavaluació del tema 3 en l'apartat b) es pot vore:  $d_2=2.326$  perquè  $n=5$ , aleshores, estimem la desviació típica com  $34.286/2.326 = 14.74$ .

#### Qüestió: **Gràfica P i U**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

Com es sap que punts estan fora de control en les gràfiques U i P? En X i R es veu clarament mirant que no estan entre LSC i LIC, però no veig com es trau per exemple en l'examen 2/2/08, ex 3, on es diu que el numero 10 està fora de control. Com es veu això?

També mirant si no està entre eixos dos límits, però si són proporcions, fent-la (en la U, igual). Fixa't que en eixe exercici els límits són 0 i 0.1984. El que més defectes té és la mostra número 9, que té 10 defectuosos de 50 que s'han mirat, per tant, la proporció és  $10/50=0.2$  que és més gran que 0.1984, aleshores, se n'ix per dalt. El següent en més defectes és la 7, que té 7 defectuosos de 50, o siga  $7/50=0.14$ , que ja no se n'ix, per tant, sols s'eixia 1 punt, el que tenia 10 defectuosos de 50 que s'han mirat.

En la pàgina 96 i 97 del llibre que us deixí a l'aulavirtual el tens resolt pas a pas, amb aquesta explicació. En la pàgina 99 i 100, tens altre exemple d'un examen però per a la gràfica U.

**Qüestió: Gráficos de control: m y n**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** En el ejemplo 3.3 que pertenece a la grafica X-bar dice que  $n=3$  y  $m=20$ , donde  $m$  es el nº de muestras, y en el ejemplo 3.4, que pertenece a la grafica de P, tengo apuntado que  $m=25$  y  $n=100$ , ¿no sería  $m=100$  porque 100 es el numero de muestras como ocurre en el ejemplo 3.3?, muchas gracias!!

Lo tienes bien apuntado,  $m=25$  y  $n=100$  en el ejemplo 3.4. La  $n$  hace referencia a cuantas muestras (muestrecillas) se toman en cada momento. En este caso se miran 100 cada vez. La  $m$  hace referencia a cuántas veces a lo largo del tiempo hemos tomado muestrecillas, que en este caso fue 25. Fíjate que dice que se toman 100 fusibles cada hora y se miran cuántos son defectuosos, por tanto,  $n=100$ . Y en total se recogen las observaciones de 25 horas, así que  $m=25$ .

**Qüestió: ICP e ICP\_k**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

Tengo una duda con los índices de capacidad. En las soluciones de exámenes anteriores he visto que cuando se analiza el ICP e ICPk lo primero que se realiza es la comparación de ambos para ver si el proceso está centrado. Bien, mi duda es, que si son diferentes y significa que el proceso no está centrado, ¿porqué el siguiente paso es comparar si es mayor o menor (el ICP) con 1.33 para comprobar la capacidad del proceso? ¿No habría que comparar el ICPk, puesto que el ICP sólo deberíamos contrastarlo en caso de que el proceso estuviera centrado?

El ICP mira la capacidad potencial del proceso, o sea, en caso de que el proceso estuviera centrado, cuál sería su capacidad.

Si  $IPC=ICP_k$  -> proceso centrado, y ambos, ICP e  $ICP_k$  son equivalentes. También podemos ver si está centrado, si la media del proceso coincide con la dimensión nominal (la media de LSE y LIE). Los dos los comparamos con 1.33 para ver si es capaz el proceso, en este caso, ambos coincidirían.

Si  $ICP \neq ICP_k$  -> proceso no centrado. El  $ICP_k$  nos daría la capacidad del proceso tal cuál está, o sea, no centrado, y lo comparamos con 1.33 en general.

El ICP nos daría la capacidad potencial, o sea, si consiguiéramos centrar el proceso, cuál sería su capacidad entonces. También lo miramos con 1.33.

El ICP será mayor que el  $ICP_k$  por cómo se obtiene. Si ya  $ICP < 1.33$  -> proceso no capaz, ¡incluso aunque estuviera centrado!

Ambos nos dan información útil.

**Qüestió: Librería qcc**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

¿Cómo se cargaba la librería de control de calidad de R?

Con library(qcc), aparece en el enunciado de la memoria. Pero debes tenerla instalada. Para ello, tienes que bajarte la librería de la web de R, <http://cran.r-project.org/web/packages/qcc/index.html>. Si estás en windows lo puedes hacer en directo desde el R, en paquetes, instalar paquetes (selecciona el mirror desde el que bajarlo) y la librería qcc. Si estás en linux, dentro de R:

```
install.packages("qcc", dependencies = TRUE)
```

En cualquier caso, lo que hará es crearte un directorio (llamado qcc), dentro del directorio library, donde tengas instalado el R.

En [http://wiki.r-project.org/rwiki/doku.php?id=getting-started:installation:packages#how\\_to\\_install\\_an\\_r\\_package](http://wiki.r-project.org/rwiki/doku.php?id=getting-started:installation:packages#how_to_install_an_r_package) tienes más explicaciones.

**Qüestió: Probabilidad de salirse del gráfico de la media**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

El apartado d del ejercicio 3 del examen 21/1/09, no lo veo claro.

El apartado d dice:

Es importante detectar rápidamente en el gráfico  $\bar{X}$ , un desplazamiento de la concentración media a 5.25 ppm, ¿cuál es la probabilidad de que el desplazamiento se detecte en la primera muestra después del corrimiento?

Así que ahora estamos ante un gráfico de la media  $\bar{X}$  y tenemos que calcular la probabilidad de que se salga.

Fíjate que cada observación de ese gráfico está formada por la media de 5 muestras, así que la variable de interés es la variable media,  $\bar{X}$  que se distribuye como una normal de media 5.25 (que dice el enunciado, si por ejemplo hubiera dicho que bajara en 0.1 respecto de la anterior, pues sería  $5.15 - 0.1 = 5.05$ , depende de que nos digan en el enunciado que queremos detectar), y desviación típica  $0.49978/\sqrt{5} = 0.2235$ . Fíjate que a diferencia del apartado c donde sólo había 1 medición, ahora estamos haciendo la media de 5, y eso se refleja en la desviación típica diviendo por la raíz del número de observaciones ( $n$ ) con las que hacemos la media.

Ahora es el momento de sacar la probabilidad de que se salga por arriba del límite superior o bien por debajo del límite inferior:

$$P(\bar{X} < 4.479) + P(\bar{X} > 5.82) = \text{Tipífico (resto la media y divido por la desviación típica)} = P(Z < (4.479-5.25)/0.2235) + P(Z > (5.82-5.25)/0.2235) = P(Z < -3.45) + P(Z > 2.55) = \text{Miramos la tabla de la Normal} = 0.0003 + 0.0054 = 0.0057$$

Fíjate que:

$$P(Z < -3.45) = (\text{segunda chuletila de la tabla}) = 1 - P(Z < 3.45) = 1 - 0.9997 = 0.0003$$

$$P(Z > 2.55) = 1 - P(Z < 2.55) = 1 - 0.9946 = 0.0054$$

En la autoevaluación del tema 3 (ejercicio 2) tienes otro problema completamente resuelto sobre esto, por si quieres ver más. También en el libro que os dejé en el aulavirtual hay de exámenes de otros años.

**Qüestió: Probabilidad de salirse del gráfico de la media y tabla de la Normal(0,1)**

(Darrera edició: divendres, 17 setembre 2010, 13:07)



**Resposta:**

En l'examen 24/01/07, exercici 1d, com es fa el càlcul de  $P(X > 0.242) + P(X < 0.1365)$ ? Sé que s'ha de mirar la taula Normal, però no sé com exactament.

El primer és determinar quina Normal hem de gastar. Com ens diu que ens fem en la situació on la mitjana augmenta en 0.04 segons i que traguem la probabilitat de detectar el canvi, és a dir, que el punt se n'isca del gràfic de la X (mitjana), aleshores, la Normal tindrà de mitjana, la d'abans + 0.04, o siga, mitjana=0.1893+0.04=0.2293. Com estem traçant la probabilitat de que se n'isca del gràfic X i cada punt que representem en eixe gràfic està format per la mitjana de 4 observacions, aleshores la desviació serà la d'abans però dividida per l'arrel de 4, o siga,  $0.035/2=0.0175$ .

Per tant, la variable mitja X és normal (0.2293, 0.0175)

Es pot eixir el punt següent per dalt o per baix dels límits del gràfic de la mitjana que havíem tret en l'apartat b), i així detectaríem que s'ha produït el canvi, aleshores hem de traure  $P(X > 0.242) + P(X < 0.1365)$ .

Per traure la probabilitat d'una Normal, primer tipifiquem i la convertim en una normal estandar (Z), o siga, restem la mitja de la normal i dividim per la desviació típica, i així podrem mirar la taula. Per tant,

$$P(X > 0.242) + P(X < 0.1365) = P\left(Z > \frac{0.242 - 0.2293}{0.0175}\right) + P\left(Z < \frac{0.1365 - 0.2293}{0.0175}\right) = P(Z > 0.73) + P(Z < -5.3)$$

Ara has de mirar la taula, però torna probabilitats de menor que un valor, és a dir,  $P(Z < x)$ .

Per tant, gasta la primera "xulla" que fica a la taula per  $P(Z > 0.73) = 1 - P(Z < 0.73) = 1 - 0.7673 = 0.2327$

Per a  $P(Z < -5.3)$  gasta la segona xulla:  $P(Z < -5.3) = 1 - P(Z < 5.3) = 1 - 1 = 0$   
5.3 ja no apareix a la taula, perquè de 3.69 endavant ja es suposa com a 1 la probabilitat.

En definitiva,  $P(Z > 0.73) + P(Z < -5.3) = 0.2327 + 0 = 0.2327$  (no ix exacte el resultat, perquè ahí vaig gastar més decimals).

En l'exercici 2 e) de l'autoavaluació del tema 3, tens altre resultat. També en la pàgina 95 del llibre que us vaig deixar a maig en l'aulavirtual.

**Qüestió: Probabilidad de ser defectuoso**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

El apartado c del ejercicio 3 del examen 21/1/09, no lo veo claro.

El apartado c dice:

Si se considera peligrosa una concentración superior a 6 ppm, ¿cuál es la probabilidad de que una medición del agua de la fuente sea considerada peligrosa?

Primero la variable que tenemos que considerar es X=concentración de sustancia tóxica que es Normal de media 5.15 y desviación típica 0.49978 según se ha obtenido en el apartado b).

Es peligrosa si supera 6, por ello tenemos que sacar la probabilidad de que supere 6, o sea, de que sea mayor que 6:  $P(X > 6)$ , esa probabilidad se saca, primero tipificando, o sea, restando la media y dividiendo por la desviación típica, como ocurre en la solución, y luego mirando la tabla de la Normal.

En la autoevaluación del tema 3 (ejercicio 2) tienes otro problema completamente resuelto sobre esto, por si quieres ver más. También en el libro que os dejé en el aulavirtual hay de exámenes de otros años.

**Qüestió: Probabilidad de ser defectuoso (2)**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En el apartado d) del ejercicio 1 de los problemas del tema 3. ¿Cómo se determina el porcentaje de elementos defectuosos en la fabricación?

Si  $X$  es la variable de interés que se distribuye Normalmente con media 431 y  $\sigma = 22.1$ , tienes que sacar la probabilidad de que esté fuera de las especificaciones, por debajo de  $430-30=400$ , o por arriba de  $430+30= 460$ , o sea,  $P(X<400)+ P(X>460)$ .

Para sacarlas tienes que usar lo que viste en IG12: tipificar y mirar la tabla de la  $Z$ , Normal  $(0,1)$  [puedes consultar en otras dudas de este mismo tema, para ampliar este punto].

Puedes mirar también las soluciones de la autoevaluación del tema 3 (ejercicio 1, apartado c), donde tienes uno completamente resuelto, o bien, el problema 3.2 de la página 85 que hicimos en clase, o la página 15 y 16 de los problemas completamente resueltos del tema 3. En el libro también encontrarás más resueltos (páginas 93 y 94).

**Qüestió: Probabilitat d'una Normal**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

En l'autoavaluació del tema 3, en l'exercici 2, en l'apartat c, per què es passa de  $P(x < 180)$  a  $P(X-223/14.74 < 180-223/14.74)$ ?

Per poder traure la probabilitat d'una Normal primer tipifiquem i la convertim en una normal estandar ( $Z$ ), o siga, restem la mitjana de la normal i dividim per la desviació típica, i així podrem mirar la taula de la Normal estandar ( $Z$ , que té mitjana 0 i desviació típica 1). Això és de l'assignatura prèvia IG12. Si no tipificàrem hauriem de tindre infinites taules per totes les possibles Normals.

En este cas, la mitjana és 223 (treta en l'apartat b)) i la desviació típica 14.74 (també treta en l'apartat b)).

Fixa't en la diferència entre l'apartat c) i e). En c) tenim només la variable de qualitat (una), però en e) és la mitjana de les 5 variables de qualitat, perquè en el gràfic de la mitjana es representen  $n$  ( $n=5$  en aquest problema) observacions cada volta.

**Qüestió: ¿Qué gráfico usar?**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

Los ejercicios 6 y 7 del tema 3 no sé cómo hacerlos, no tengo ni idea.

El 6 es un gráfico  $P$ , pues estamos con una Binomial y miramos proporción de defectuosos. Es como el problema 3.4 de los apuntes hecho en clase. También puedes mirar la página 789 de problemas resueltos de Montgomery, donde verás cómo calcular la gráfica  $P$ .

En el 7, hay que ver que es Poisson y por tanto usaríamos un gráfico  $U$  para controlar el número de camiones que llegan a) en una hora, b) en un día. Sólo hay que sustituir en la fórmula del gráfico  $U$ , con  $n=1$ , pues sólo miramos uno (hora o día según el apartado) cada vez.

En el apartado 4.4. del libro del aulavirtual tienes problemas resueltos para gráficos de control de atributos.

## TEMA 4

### Qüestió: **Datos de los problemas del tema 4**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

#### **Resposta:**

He encontrado el fichero dades problemes tema 4 (probt4.txt) en el aulavirtual pero al abrirlo no hay valores en las variables Entrada/Salida, Longitudcola, Disciplina y Entorno. ¿Me puede ayudar?

Sí que están ahí los valores, sólo que al estar separados por tabuladores, según con el editor de texto con que lo abras, no verás los datos justo debajo de la columna respectiva. Guarda el fichero y léelo con el R, como en prácticas (mira por ejemplo la práctica 4).

### Qüestió: **Grup homogeni**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

#### **Resposta:**

En el tema 4, què vol dir grups homogenis? Per exemple, en el problema 3 de l'examen de setembre de 2007.

Un grup homogeni és aquell en el que els seus elements tenen la mateixa mitjana poblacional.

Has de fer la comparació de totes les parelles, en aquest cas com hi ha 5 grup, tens 10 comparacions, que és el que apareix en la solució. Si tingueres 4 grups, serien 6 comparacions.

En aquest problema hi ha dos grups homogenis: primer, el format pel 1 i 2 (4000 i 3300AC), i el segon pel grup del 3, 4 i 5 (1850AC, 200AC, 150DC), que és el que es pot vore en la solució. El 1 i 2 tenen igual mitjana, però distinta de la del 3, 4 i 5. El 3, 4 i 5 tenen iguals mitjanes. La X en 1 i 2 estan en la mateixa columna, la X en 3, 4 i 5 estan en la mateixa columna, però distinta de la del 1 i 2.

Has d'ordenar les mitjanes mostrals de menor a major i col·locar una X en distinta columna si existeix diferència entre les mitjanes poblacionals.

### Qüestió: **LSD ANOVA de 1 vía**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** No me ha quedado claro qué son la  $n_{subi}$  y la  $n_{subj}$  en la fórmula de la LSD en la página 112 del libro (diseño completamente aleatorizado o análisis de la varianza con un factor).

$n_i$  y  $n_j$  son el número de observaciones que hay en cada grupo, en el grupo  $i$  y grupo  $j$ , que podría ser distinto y tendríamos que calcular distintas LSD para cada comparación de grupos, pero... para simplificar los cálculos, os dije que en los exámenes, todos los grupos tendrían el mismo número de observaciones, por tanto, en el caso particular de nuestros exámenes,  $n_i$  y  $n_j$  SIEMPRE serán iguales, igual al número de observaciones que hay en cada grupo. De hecho, en los apuntes del tema 4 (página 101) verás que aparecía la señal de curvas peligrosas para que tuvieráis cuidado con ello.

En el examen de enero de 2009 (ejemplo 5.3 del libro, página 113), tenemos 4 políticas y 5 observaciones para cada una de ellas. En ese ejemplo,  $n_i = n_j = 5$ .

Puedes apuntartelo al formulario para no olvidarte.

### Qüestió: **LSD con el R**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:**

Tengo una duda sobre la práctica 4. Se trata en los apartados en los que se pide usar el método de la LSD de Fisher. En principio utilizamos la función multcomp de ANOVA, que permite hacer comparaciones múltiples y de ahí nos quedamos con la primera tabla (\$table) que es la que nos interesa. La duda me surge a partir de ahí, cuando utilizamos la función model.tables con la que calculamos las medias y a partir de ahí sacamos el gráfico de las X's. Pero... ¿con qué se compara cada par de medias? En los ejercicios de clase tengo claro que se compara con LSD al nivel alfa que calculamos mediante una fórmula dada, pero ¿y en la práctica dónde se calcula la LSD?

Con model.tables sólo se saca las medias muestrales, nada más, es sólo para ordenar las medias de menor a mayor en la tabla de las X. Para después poner las X, nos tenemos que basar en la \$table.

Coge por ejemplo, el ejercicio 4.3 de los apuntes, en clase calculamos la  $LSD=0.4689$  y vimos que es equivalente a cómo lo hace el R. Fíjate que la diferencia entre la media del grupo 1 y 2 es  $6.793-7.535 = -0.75$  y  $-0.75 \pm 0.4689$  (la LSD), es justamente el intervalo del R:  $-1.2214, -0.2835$ .

Por ello, con la salida del R, miramos si el cero pertenece o no al intervalo correspondiente (si cero está fuera, las medias son distintas). Esto es equivalente a lo que hacemos en la clase de teoría, sacar la diferencia en valor absoluto y ver si supera o no la LSD.

El R ha calculado la LSD para construir el intervalo. Si quieres saber la LSD sólo tienes que hacer  $-0.75 - (-1.2214) = 0.4689$ .

**Qüestió: Suma de cuadrados**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** Estoy con las tablas ANOVA y ahora mismo no recuerdo como se sacan la suma de cuadrados de tratamiento, de error y de total, porque las formulas no los acabo de entender, ¿me podrías poner un ejemplo en numeros de como se hallan estas formulas, por ejemplo con los datos del ejemplo 4.2.?

Por ejemplo en la suma de cuadrados total se tendría que sacar primero la media de TODAS las observaciones, en este caso hay 15 (sería la  $\bar{y}$ ). Luego, tendrías que hacer el cuadrado de la resta de cada observación menos esa media, y sumar todos esos resultados para las 15 observaciones. Fíjate que sería como sacar a varianza de todos los datos, pero sin dividir por  $N-1$ . Con la calculadora, podrías sacar la desviación típica de los 15 datos, hacer el cuadrado para sacar la varianza, y luego multiplicarlo por  $N-1$  para tener sólo la suma pedida.

Las otras sumas de cuadrados son análogas, pero donde  $\bar{y}_i$  es la media de cada grupo.

Como hacer todos esos cálculos manualmente es laborioso, y con el ordenador ya vimos que está hecho rápidamente, os dije que en el examen os pondría ya parte de la tabla calculada y sólo tendríais que acabar de rellenarla, conociendo las relaciones entre las celdas de la tabla.

En el libro del aulavirtual <https://aulavirtual.uji.es/mod/resource/view.php?id=334200>, tienes ejemplos completamente resueltos.

**Qüestió: X's en LSD**

(Darrera edició: divendres, 17 setembre 2010, 13:07)

**Resposta:** En el problema 2 b) ii del examen de 21/1/2009, cuando pides los grupos homogéneos, no entiendo muy bien dónde poner las Xs.

Para colocar las X, primero habrás visto con la LSD dónde, entre qué grupos hay o no diferencias de medias. Después ordena los grupos de menor a mayor media, y vas colocando una X en una columna distinta de la anterior si hay diferencia, si no hay diferencia la X la colocas a la misma altura (en la misma columna que el grupo anterior). Justo ese problema está resuelto en el libro que os dejé en el aulavirtual.

Así:

Política

2 X

3

1

4

Después, la 2 y 3 no difieren, por tanto, en la misma columna

2 X

3 X

1

4

Después miro la 1 y la 2 que sí difieren, pero la 1 y 3 no difieren. La única forma de expresar eso, es meter 2 X's al 3, una X a la altura del 2 (que ya estaba) y otra en una columna distinta que estará con el 1. Esta última está en una columna distinta de la del 2, lo cual indica que hay diferencia entre la media del 1 y 2, justo lo que queríamos expresar. B indica un espacio en blanco.

2 X

3 XX

1 BX

4

Por último, el 4 difiere de todas así que la X va en columna aparte. B indica un espacio en blanco.

2 X

3 XX

1 BX

4 BBX

Amb més detall, encara:

Anem de dalt cap baix mirant cada grup amb els anteriors, amb aquesta metodologia es poden resumir totes les comparacions.

La primera no es compara amb ningú, de moment sols està el grup 2. Per això, fiquem una X, i ja està.

Política

2 X

3

1

4

Després anem al següent que és el grup 3, que sols podem comparar amb el 2 que és l'únic que ja està clar. Com són iguals les mitjanes, a la mateixa columna la X:

2 X

3 X

1

4

El 3 no el mirem ni amb 1 ni amb 4 de moment, ja es faran quan els toque el torn.

Després anem al següent que és el grup 1, que compararem amb els anteriors. Primer 2 i 1. 1 i 2 diferixen, per tant, X en columna apart (B indica espai en blanc)

2 X  
3 X  
1 BX  
4

Ara mire el 1 i 3, 1 i 3 no diferixen, per tant, li fiqui altra X al 3 a la mateixa columna que l'1, per indicar que les dos són iguals.

2 X  
3 XX  
1 BX  
4

El 4 és diferent de totes, la X va en columna apart (miraria 2-4 primer i veig que he de ficar-la en columna apart, miraria després que 3-4 diferixen, i per últim la 1-4 que com també diferixen aniria la X en columna apart):

2 X  
3 XX  
1 BX  
4 BBX

La lectura és la següent: si la X està a la mateixa columna, les mitjanes NO diferixen, si estan en columnes diferents, són distintes les mitjanes.