Examen Final Problemas, Enero 2015

Supongamos que en el examen teórico que acabamos de realizar, que ha constado de 20 preguntas y cada una de ellas con 4 posibles respuestas, por exámenes anteriores se sabe que la probabilidad de contestar correctamente a cualquier pregunta es 0.65. El examen se puntúa un punto positivo, si la respuesta es correcta y medio punto negativo si es incorrecta o en blanco. Aquellos alumnos que contesten bien más de 13 preguntas superan el examen.

1. **(1) ¿Probabilidad de que un alumno elegido al azar supere el examen?.**

2. **(1) Si en el curso hay matriculados 150 alumnos, ¿cuál es la probabilidad de que superen el examen como mínimo el 50% de los matriculados?.**

Del total de alumnos del curso, el 40% son alumnos repetidores, que por experiencias anteriores se sabe que superan el examen con probabilidad 0.75,

3. **(1) Elegido un alumno al azar, ha resultado que está aprobado, ¿probabilidad de que este alumno sea de 1ª matricula?.**

4. **(1) Obtener la Nota Media del examen (sobre 10), así como su Desviación Típica, (sobre 20).**

El profesorado del Departamento de Estadística afirma que las notas de Teoría son superiores a las notas de Problemas. Para comprobar tal afirmación se toma una muestra de 15 alumnos (m.a.s.) y observa las notas (sobre 10) en el examen Teórico y de Problemas. Los resultados son:

Notas Teoría	**5.5**	**5**	**7**	**6**	**4**	**3.8**	**6.2**	**5**	**5.1**	**5**	**4.7**	**4.8**	**5**	**8**	**5.1**
Notas Problemas	**5**	**5**	**4**	**3**	**2**	**6**	**5.2**	**4.9**	**4.8**	**3.7**	**2**	**3**	**4**	**5.1**	**6.2**

Se proporcionan los siguientes datos auxiliares:

$$\sum T_i = 80.2 \;;\; \sum T_i^2 / 15 = 29.685\hat{3} \;;\; \sum P_i = 63.9 \;;\; \sum P_i^2 = 296.23 \;;\; \sum T_i P_i = 345.24$$

5. **(1) ¿Es cierta tal afirmación?.**

6. **(0.5) ¿Cuál es la probabilidad límite del contraste anterior?.**

7. **(0.5) ¿Cuál sería el tamaño mínimo necesario para estimar la media en el examen de Teoría con una precisión de ±0.5 y un coeficiente de confianza del 99%?.**

Suponemos que las Horas de estudio de cada alumno sigue una v.a. cuya función de densidad es:

$$p(x) = \frac{\lambda^x}{(1+\lambda)^{1+x}} \quad X = 0, 1, 2, \ldots .$$

de la que se sabe que: $E(X) = \lambda \;;\; E(X^2) = \lambda + 2\lambda^2$

8. **(1.5) ¿El estimador por máxima verosimilitud del parámetro λ, es consistente?.**

9. **(0.5) ¿Cuánto vale el grado de relación entre las Notas de Teoría y de Problemas?.**

10. **(1) Si un alumno ha obtenido 6.5 en el examen de Problemas, ¿Qué nota habrá sacado en el examen de Teoría?.**

11. **(1) Sin considerar las Notas numéricas, solamente si un alumno Aprueba o Suspende, tanto en Teoría como en Problemas, ¿cuál es el grado de relación entre ambas Notas, si es que existe?.**

SOLUCIÓN

1. **¿Probabilidad de que un alumno elegido al azar supere el examen?. (Tomar 4 decimales).**

Sea A_i la v.a. que mide si una pregunta está bien respondida

$$A_i \in B(p = 0.65)$$

Sea la v.a. A≡ "Número de preguntas correctamente contestadas"

$$A = \sum_{i=1}^{20} A_i \in b(n = 20; p = 0.65)$$

Se supera el examen si A>13

$$P(A > 13) = \sum_{a=14}^{20} \binom{20}{a} \times 0.65^a \times 0.35^{20-a} = \binom{20}{14} \times 0.65^{14} \times 0.35^6 + ... + \binom{20}{20} \times 0.65^{20} \times 0.35^0 =$$

$$= 0.1712 + 0.1272 + 0.0738 + 0.0322 + 9.984 \times 10^{-3} + 1.951 \times 10^{-3} + 1.812 \times 10^{-4} = \underline{0.4175}$$

2. **Si en el curso hay matriculados 150 alumnos, ¿cuál es la probabilidad de que superen el examen como mínimo el 50% de los matriculados?.**

Un alumno supera el examen con probabilidad 0.4175

$$S_i \in B(p = 0.4175)$$

En el conjunto de los 150 alumnos del curso, la v.a. S que mide cuantos alumnos han superado el examen, tendrá por distribución:

$$S = \sum_{i=1}^{150} S_i \in b(n = 150; p = 0.4175) \xrightarrow{\text{T.C.L.}\,(n \geq 30;\ \min(p,q) \geq 0.1)}$$

$$S^* \in N(\mu = 150 \times 0.4175\ ;\ \sigma^2 = 150 \times 0.4175 \times 0.5825) \equiv N(\mu = 62.625\ ;\ \sigma^2 = 36.479)$$

Si han de superar el examen como mínimo el 50% de los alumnos matriculados, $S \geq 75$, con probabilidad:

$$P(S \geq 75) \approx P(S^* \geq 75 - 0.5) = P\left(Z \geq \frac{74.5 - 62.625}{\sqrt{36.479}}\right) = P(Z \geq 1.9661) \approx \underline{0.0244}$$

Sin aplicar Yates: (no válido)

$$P(S \geq 75) \approx P(S^* \geq 75) = P\left(Z \geq \frac{75 - 62.625}{\sqrt{36.479}}\right) = P(Z \geq 2.0489) \approx \underline{0.0202}$$

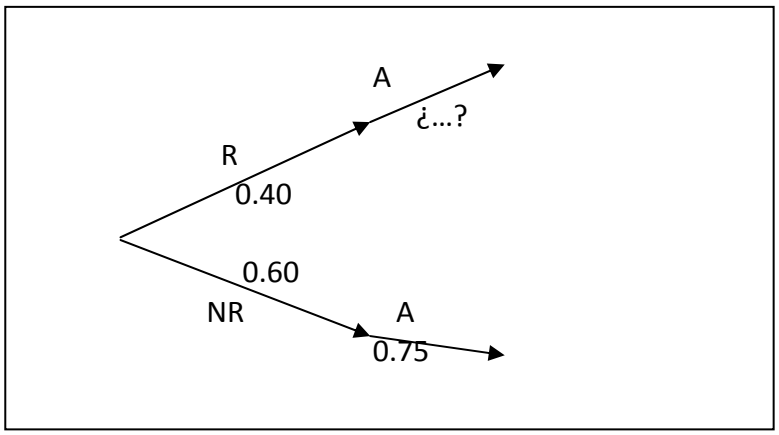
Del total de alumnos del curso, el 40% son alumnos repetidores, que por experiencias anteriores se sabe que superan el examen con probabilidad 0.75,

3. **Elegido un alumno al azar, ha resultado que está aprobado, ¿probabilidad de que este alumno sea de 1ª matricula?.**

Sean los sucesos: R ≡ "Alumno repetidor"
NR ≡ "Alumno de 1ª matricula"
A ≡ "Aprobar el examen"

Según el enunciado: $P(R)=0.40 \Rightarrow P(NR)=0.60$

$$P\left(A/_{R}\right)=0.75\ ;\quad P\left(A/_{NR}\right)=¿...?\quad P(A)=0.4175$$



De acuerdo con el Teorema de la Probabilidad Total:

$$P(A)=P(R)\times P\left(A/_{R}\right)+P(NR)\times P\left(A/_{NR}\right)$$

$$0.4175=0.40\times 0.75+0.60\times P\left(A/_{NR}\right)\Rightarrow P\left(A/_{NR}\right)=0.1958$$

Aplicando el Teorema de Bayes:

$$P\left(NR/_{A}\right)=\frac{P(NR\cap A)}{P(A)}=\frac{P(NR)\times P\left(A/_{NR}\right)}{P(R)\times P\left(A/_{R}\right)+P(NR)\times P\left(A/_{NR}\right)}$$

Sustituyendo valores:

$$P\left(NR/_{A}\right)=\frac{0.60\times 0.1958}{0.4175}=\underline{0.2813}$$

4. **Obtener la Nota Media del examen (sobre 10), así como su Desviación Típica (sobre 20).**

Nota que se obtiene en el examen, sobre 20:

$$N=1\times A+(-0.5)\times(20-A)=1.5A-10$$

Siendo A, el Número de preguntas correctamente contestadas.

Recordemos que:

$$A \in b(n=20; p=0.65) \Rightarrow \begin{cases} \mu_A = E[A] = 20 \times 0.65 = 13 \\ \sigma_A^2 = V(A) = 20 \times 0.65 \times .035 = 4.55 \end{cases}$$

- ✓ Nota Media (sobre 20): $\mu_N = E[N] = E[1.5A - 10] = 1.5 \times E[A] - 10$

sustituyendo:

$$\mu_N / 20 = \{1.5 \times E[A] - 10\} / 20 = (1.5 \times 13 - 10) / 20 = \underline{9.5/20}$$

- ✓ Nota Media (sobre 10): $\Rightarrow \begin{Bmatrix} 20 - 10 \\ 9.5 - \overline{N} \end{Bmatrix} \Rightarrow \overline{N} = \frac{10 \times 9.5}{20} = \underline{4.75}$

- ✓ Desviación Típica (sobre 20): $\sigma_N = +\sqrt{\sigma_N^2}$

$$\sigma_N^2 / 20 = V(N) / 20 = V(1.5A - 10) / 20 = [1.5^2 \times V(A)] / 20$$

sustituyendo:

$$\sigma_N^2 / 20 = [1.5^2 \times V(A)] / 20 = (2.25 \times 4.55) / 20 = 10.2375/20 \Rightarrow \sigma_N / 20 = \sqrt{10.2375} / 20 = \underline{3.1996/20}$$

El profesorado del Departamento de Estadística afirma que las notas de Teoría son superiores a las notas de Problemas, para comprobar tal afirmación toma una muestra de 15 alumnos (m.a.s.) y observa las notas (sobre 10) en el examen Teórico y de Problemas, los resultados son:

Notas Teoría	**5.5**	**5**	**7**	**6**	**4**	**3.8**	**6.2**	**5**	**5.1**	**5**	**4.7**	**4.8**	**5**	**8**	**5.1**
Notas Problemas	**5**	**5**	**4**	**3**	**2**	**6**	**5.2**	**4.9**	**4.8**	**3.7**	**2**	**3**	**4**	**5.1**	**6.2**

5. ¿Es cierta tal afirmación?.

Se trata de muestras procedentes de poblaciones Relacionadas, (son los mismos alumnos)

Notas **T**eoría	5.5	5	7	6	4	3.8	6.2	5	5.1	5	4.7	4.8	5	8	5.1
Notas **P**roblemas	5	5	4	3	2	6	5.2	4.9	4.8	3.7	2	3	4	5.1	6.2
Diferencia=T-P	0.5	0	3	3	2	-2.2	1	0.1	0.3	1.3	2.7	1.8	1	2.9	-1.1

- ✓ Test de hipótesis a contrastar: $\begin{cases} H_0 : \mu_T \le \mu_P \\ H_1 : \mu_T > \mu_P \end{cases} \Rightarrow \begin{cases} H_0 : \mu_D \le 0 \\ H_1 : \mu_D > 0 \end{cases}$
- ✓ Estadístico del contraste: $T = \dfrac{\overline{D} - \mu_{D_0}}{S_D / \sqrt{n_D - 1}} \in t_{(n-1)}$
- ✓ Región de Aceptación si H_0 es cierta: $C_0 = \left(-\infty \, ; \, t_{\alpha;(n-1)}\right)$

Datos muestrales:

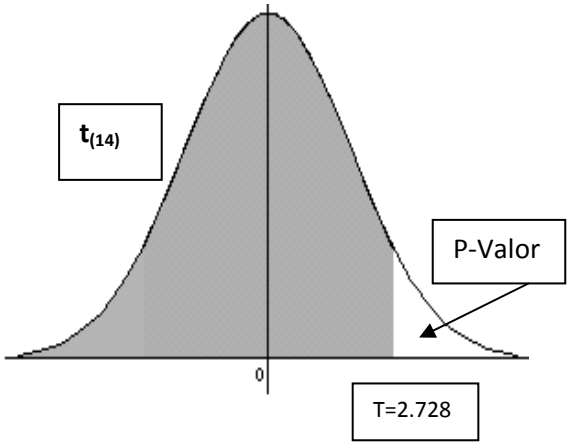
$$\sum T_i = 80.2 \; ; \; \sum T_i^2 = 445.28 \; ; \; \sum P_i = 63.9 \; ; \; \sum P_i^2 = 296.23 \; ; \; \sum T_i P_i = 345.24$$

$$\begin{cases} \overline{T} = 5.34\widehat{6} \; ; \; S_T = 1.04808 \\ \overline{P} = 4.26 \; ; \; S_P = 1.26533 \end{cases} ; \; \overline{D} = 1.08\widehat{6} \; ; \; S_D = 1.49035 \; ; \; n_D = 15$$

- ➢ $T = \dfrac{1.08\widehat{6}}{1.49035 / \sqrt{15-1}} = 2.728$
- ➢ $\alpha = 0.05 \Rightarrow t_{0.05;(14)} = 1.761$
- ➢ $C_0 = \left(-\infty \, ; \, 1.761\right)$, como $\underline{T = 2.728 \notin C_0 = \left(-\infty \, ; \, 1.761\right)}$

RECHAZAMOS H_0, es decir, la afirmación del Departamento es cierta, las Notas de Teoría son superiores a las de Problemas.

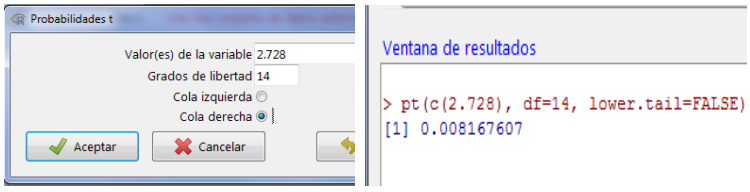
6. ¿Cuál es la probabilidad límite del contraste anterior?.



- ✓ **Según las tablas proporcionadas en Moodle:**

$$P - \text{value} = P\left(t_{(14)} > 2.728\right) \approx \underline{0.01}$$

- ✓ **Según el paquete estadístico R:**



$$P - \text{value} = P\left(t_{(14)} > 2.728\right) = 0.008167607$$

7. ¿Cuál sería el tamaño mínimo necesario para estimar la media en el examen de Teoría con una precisión de ±0.5 y un coeficiente de confianza del 99%?.

Intervalo de Confianza para la media poblacional en el examen de Teoría:

$$I_{0.99}\left(\mu_T\right)=\left(\overline{T}-t_{0.005;(14)}\frac{S_T}{\sqrt{N-1}}\;;\;\overline{T}+t_{0.005;(14)}\frac{S_T}{\sqrt{N-1}}\right)=\left(\overline{T}-\varepsilon\;;\;\overline{T}+\varepsilon\right)$$

siendo $\varepsilon=t_{0.005;(14)}\frac{S_T}{\sqrt{N-1}}$

De acuerdo con los resultados muestrales, (muestra piloto)

$$\left.\begin{cases}\varepsilon=0.5\\ S_T=1.04808\\ t_{0.005;(14)}=2.977\end{cases}\right\}\Rightarrow 0.5=2.977\times\frac{1.04808}{\sqrt{N-1}}\Rightarrow N=39.9409\;\;;\;\;\underline{N=40}$$

Suponemos que las Horas de estudio de cada alumno sigue una v.a. cuya función de densidad es:

$$\mathbf{p\left(x\right)=\frac{\lambda^{x}}{\left(1+\lambda\right)^{1+x}}\;\;X=0,1,2,....}$$

de la que se sabe que: $\mathbf{E\left(X\right)=\lambda\;\;;\;\;E\left(X^2\right)=\lambda+2\lambda^2}$

8. ¿El estimador por máxima verosimilitud del parámetro λ, es consistente?

Obtención del estimador por máxima verosimilitud para el parámetro λ.

$$p\left(x\right)=\frac{\lambda^{x}}{\left(1+\lambda\right)^{1+x}}\;\;X=0,1,2,....$$

- ✓ Función de Verosimilitud:

$$L\left(\vec{X};\lambda\right)=\prod_{i=1}^{n}p\left(x_i;\lambda\right)=\frac{\lambda^{X_1}}{\left(1+\lambda\right)^{1+X_1}}\times\frac{\lambda^{X_2}}{\left(1+\lambda\right)^{1+X_2}}\times...\times\frac{\lambda^{X_n}}{\left(1+\lambda\right)^{1+X_n}}$$

$$L\left(\vec{X};\lambda\right)=\frac{\lambda^{\sum_{i=1}^{n}X_i}}{\left(1+\lambda\right)^{\sum_{i=1}^{n}(1+X_i)}}=\frac{\lambda^{\sum_{i=1}^{n}X_i}}{\left(1+\lambda\right)^{n+\sum_{i=1}^{n}X_i}}$$

- ✓ Logaritmo Neperiano de la Función de Verosimilitud:

$$Ln\left[L\left(\vec{X};\lambda\right)\right]=\left(\sum_{i=1}^{n}X_i\right)Ln\,\lambda-\left(n+\sum_{i=1}^{n}X_i\right)Ln\left(1+\lambda\right)$$

- ✓ Derivada del logaritmo Neperiano de la Función de Verosimilitud:

$$\frac{\partial \operatorname{Ln}\left[L\left(\vec{X};\lambda\right)\right]}{\partial\lambda}=\frac{\left(\sum_{i=1}^{n}X_i\right)}{\lambda}-\frac{\left(n+\sum_{i=1}^{n}X_i\right)}{1+\lambda}$$

- ✓ Igualando a cero la derivada del logaritmo Neperiano de la Función de Verosimilitud y haciendo operaciones:

$$\frac{\left(\sum_{i=1}^{n}X_i\right)}{\lambda}-\frac{\left(n+\sum_{i=1}^{n}X_i\right)}{1+\lambda}=0$$

$$\frac{\left(\sum_{i=1}^{n}X_i\right)}{\lambda}=\frac{\left(n+\sum_{i=1}^{n}X_i\right)}{1+\lambda}\Rightarrow(1+\lambda)\left(\sum_{i=1}^{n}X_i\right)=\lambda\left(n+\sum_{i=1}^{n}X_i\right)$$

$$\sum_{i=1}^{n}X_i=\lambda n\Rightarrow\hat{\lambda}=\frac{\sum_{i=1}^{n}X_i}{n}=\overline{X}$$

¿Es consistente el estimador obtenido?

Consistente: Cuando siendo al menos asintóticamente insesgado:

$$V\left(\hat{\lambda}\right)\xrightarrow[n\to\infty]{}0$$

¿Es Insesgado?

Insesgado: $E\left(\hat{\lambda}\right)=\lambda$

$$E\left(\hat{\lambda}\right)=E\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=E(X)$$

siendo: $E(X)=\sum_{x=0}^{\infty}x\,p(x)=\sum_{x=0}^{\infty}x\,\frac{\lambda^x}{(1+\lambda)^{1+x}}$, pero del enunciado se sabe que: $E(X)=\lambda$, por tanto el estimador obtenido es insesgado.

Calculemos ahora la varianza del estimador obtenido:

$$V\left(\hat{\lambda}\right)=\frac{1}{n^2}\sum_{i=1}^{n}V(X_i)=\frac{V(X)}{n}$$

$$V(X)=\sum_{x=0}^{\infty}(x-\lambda)^2p(x)=\alpha_2-\alpha_1^2$$

siendo $\alpha_2=\sum_{x=0}^{\infty}x^2\,p(x)=\sum_{x=0}^{\infty}x^2\,\frac{\lambda^x}{(1+\lambda)^{1+x}}$, pero del enunciado se sabe que: $\alpha_2=E\left(X^2\right)=\lambda+2\lambda^2$

por tanto: $V(X) = \lambda + 2\lambda^2 - \lambda^2 = \lambda + \lambda^2 = \lambda(1+\lambda)$

$$V(\hat{\lambda}) = \frac{\lambda(1+\lambda)}{n} \xrightarrow[n\to\infty]{} 0$$

El estimador obtenido **es Consistente**

9. **¿Cuánto vale el grado de relación entre las Notas de Teoría y de Problemas?**

   Por tratarse de dos variables numéricos, calcularemos el coeficiente de correlación lineal de Pearson: $R_{TP} = \frac{S_{TP}}{S_T S_P}$

   Datos muestrales: $\begin{cases} \overline{T} = 5.34\widehat{6} \ ; \ S_T = 1.04808 \\ \overline{P} = 4.26 \ ; \ S_P = 1.26533 \end{cases}$

$$S_{TP} = \left(\frac{1}{15}\sum_{i=1}^{15} T_i P_i\right) - \overline{T}\,\overline{P} = \frac{345.24}{15} - 5.34\widehat{6} \times 4.26 = 0.2392$$

$$R_{TP} = \frac{0.2392}{1.04808 \times 1.26533} = \underline{0.18036}$$

Relación lineal creciente, positiva pero escasa.

10. **Si un alumno ha obtenido 6.5 en el examen de Problemas, ¿Qué nota habrá sacado en el examen de Teoría?.**

Se nos pide un modelo de Regresión lineal: $\hat{T} = a + b\,P$

donde: $\begin{cases} b = \frac{S_{TP}}{S_P^2} \\ a = \overline{T} - b\,\overline{P} \end{cases}$; $\begin{cases} b = \frac{0.2392}{1.26533^2} = 0.1494 \\ a = 5.34\widehat{6} - 0.1494 \times 4.26 = 4.7102 \end{cases}$

El modelo obtenido es: $\hat{T} = 4.7102 + 0.1494\ P$

La estimación pedida: $\hat{T}_{(P=6.5)} = 4.7102 + 0.1494 \times 6.5 = \underline{5.6813}$

11. **Sin considerar las Notas numéricas, solamente si un alumno Aprueba o Suspende, tanto en Teoría como en Problemas, ¿cuál es el grado de relación entre ambas Notas, si es que existe?.**

De acuerdo con la nueva nomenclatura, la tabla quedará:

Notas Teoría	5.5	5	7	6	4	3.8	6.2	5	5.1	5	4.7	4.8	5	8	5.1
Notas Problemas	5	5	4	3	2	6	5.2	4.9	4.8	3.7	2	3	4	5.1	6.2

Notas Teoría	A	A	A	A	S	S	A	A	A	A	S	S	A	A	A
Notas Problemas	A	A	S	S	S	A	A	S	S	S	S	S	S	A	A

Notas Teoría / **Notas Problemas**	**A**prueba	**S**uspende	$\mathbf{n_{i.}}$
Aprueba	5	1	6
Suspende	6	3	9
$\mathbf{n_{.j}}$	11	4	n = 15

Tabla de frecuencias esperadas en el caso de no relación: $f_{e_{ij}} = \dfrac{n_{i.} \times n_{.j}}{n}$

Notas Teoría / **Notas Problemas**	**A**prueba	**S**uspende	$\mathbf{n_{i.}}$
Aprueba	4.4	1.6	6
Suspende	6.6	2.4	9
$\mathbf{n_{.j}}$	11	4	n = 15

Como $f_{e_{ij}} \neq n_{ij}$ **las dos variables están relacionadas**.

El grado de relación entre ambas variables será:

Coeficiente de Pearson: $\chi^2 \equiv P = \sum_{\forall ij} \dfrac{\left(n_{ij} - f_{e_{ij}}\right)^2}{f_{e_{ij}}}$

$$\chi^2 \equiv P = \frac{(5-4.4)^2}{4.4} + \frac{(1-1.6)^2}{1.6} + \frac{(6-6.6)^2}{6.6} + \frac{(3-2.4)^2}{2.4} = 0.51136$$

Coeficiente V de Cramer: $V = \sqrt{\dfrac{P}{n \times t}} = \sqrt{\dfrac{0.51136}{15 \times 1}} = 0.18463$

Coeficiente C de Contingencia: $C = \sqrt{\dfrac{P}{P+n}} = \sqrt{\dfrac{0.51136}{0.51136+15}} = 0.18156$