1.2. Estudiar casos típicos de distribuciones estadísticas de probabilidades

Objetivos:

1. Trabajar las distribuciones estadísticas a partir de las frecuencias relativas (probabilidades).
2. Averiguar la distribución probabilística del número de ocurrencias de un suceso entre un número de ocasiones independientes. (distribución binomial).
3. Aproximar la distribución probabilística del número de ocurrencias de un suceso raro conociendo el número medio de ocurrencias entre un número grande de ocasiones independientes (distribución de Poisson).
4. Introducir la distribución de densidad probabilística de una variable aleatoria que varía de forma continua.
5. Estudiar la distribución de densidad probabilística de la media de un gran número de variables aleatorias equivalentes independientes (distribución normal).

Actividad 1.11. Para comparar distribuciones estadísticas sobre diferentes poblaciones deberíamos utilizar las correspondientes frecuencias relativas o probabilidades , definidas por $p(x)=f(x)/n(U)$.
Demostrar los siguientes teoremas:
Teorema 1.10: $\sum_x p(x) = 1$ (llamaremos *distribución probabilística* a cualquier aplicación $p:V\to R^+ +\{0\}$ que cumpla esta propiedad, siendo V un conjunto numerable de valores .
Teorema 1.11: $\mu(X) = \sum_x x\cdot p(x)$ (utilizaremos esta expresión para definir la media de cualquier distribución probabilísitica con independencia del tamaño finito o infinito de la población .
Ejercicio 1.4: representar gráficamente en diagramas de tarta las distribuciones probabilísticas de las variables aleatorias de la actividad 1.3.

Actividad 1.12. Si tenemos un conjunto A de valores de una variable aleatoria, su probabilidad vendrá definida por
$p(A) = \sum_{x\in A} p(x)$. Demostrar el
Teorema 1.12: Si A y B son dos conjuntos disjuntos de valores de una variable aleatoria, $p(A+B) = p(A) + p(B)$.

Actividad 1.13. Diremos que dos variables X, Y son independientes si para cualquier par (x, y) de valores respectivos de las mismas se cumple $p(x,y)=p(x)\cdot p(y)$.
Problema 1.6: estudiar si las variables aleatorias de la Actividad 1.3 son independientes.

Actividad 1.14. Teniendo en cuenta el
Teorema -1.1: el número de maneras en que podemos escoger m elementos entre n es

$$\binom{n}{m} = \frac{n(n-1)(n-2)...(n-(m-1))}{m!} = \frac{n!}{m!(n-m)!}$$ (*combinaciones* de n sobre m)

demostrar el
Teorema 1.13: Si tenemos n variables-ocasiones independientes con un determinado valor-suceso con la misma probabilidad p, la probabilidad de no ocurrencia de cada

suceso será q=1-p y la probabilidad de que el número de ocurrencias del suceso sea exactamente m será

$P^B(m) = \binom{n}{m} p^m q^{n-m}$ (*distribución binomial* B(p,n))

Para demostrarlo, estudiar primero la probabilidad de una determinada serie ordenada de m ocurrencias y n-m no ocurrencias, y después el número de maneras de ordenar m ocurrencias y n-m no ocurrencias, teniendo en cuenta que son indiferentes las permutaciones entre sí de las ocurrencias y las no ocurrencias.

Problema 1.7: calcular la probabilidad de obtener exactamente 3 ases en 5 lanzamientos de un dado.

Actividad 1.15. Teniendo en cuenta el

Teorema -1.2: $(a+b)^n = \sum_{m=0}^{n} \binom{n}{m} a^m b^{n-m}$ (*binomio de Newton*)

demostrar el

Teorema 1.14: $\sum_{m=0}^{n} P^B(m) = 1$.

Actividad 1.16. Teniendo en cuenta el

Teorema -1.3: $0!=1$, $m!=m\cdot(m-1)!$

demostrar los

Teorema 1.15: para todo m=1...n, $\binom{n}{m}\cdot m = n\cdot\binom{n-1}{m-1}$

Teorema 1.16: $\mu(B(p,n)) = np$.

Actividad 1.17. Demostrar los

Teorema 1.17: para todo m=2...n, $m\cdot\binom{n-1}{m-1} =)n-1)\cdot\binom{n-2}{m-2} + \binom{n-1}{m-1}$

Teorema 1.18: $\sigma(B(p,n))^2 = npq = np(1-p)$

Actividad 1.18.

Problema 1.8: obtener la media y la desviación típica del número de ases al lanzar 30 veces un dado.

Actividad 1.19. Si p es muy pequeño, para obtener una media μ apreciable de ocurrencias de un suceso necesitaremos un número n muy grande de ocasiones. Pero los factoriales n!, y por lo tanto la distribución binomial, son difíciles de calcular si n es grande. En este caso, habremos de utilizar una aproximación. A tal efecto, y teniendo en cuenta que $e = \lim_{u\to\infty} (1+1/u)$, y por lo tanto

Teorema -1.4: $\lim_{n\to\infty} (1-\mu/n) = e^{-\mu}$

demostrar el

Teorema 1.19: si $p=\mu/n$, $P^\Pi(m) = \lim_{n\to\infty} P^B(m) = e^{-\mu}\cdot\mu^m/m!$ (*distribución de Poisson* $\Pi(\mu)$).

La distribución de Poisson es una buena aproximación a la binomial si n>50, p<0'1 y μ=np<5 .

Actividad 1.20. Teniendo en cuenta el

Teorema -1.5: $e^{\mu} = \sum_{m=0.}^{\infty} \mu^m/m!$ (desarrollo en serie de Taylor del exponencial)

demostrar los

Teorema 1.20: $\sum_{m=0.}^{\infty} P^{\Pi}(m) = 1.$

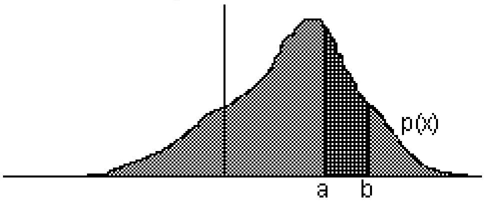
Teorema 1.21: $\mu(\Pi(\mu)) = \mu$

Teorema 1.22: $\sigma(\Pi(\mu))^2 = \mu$

Actividad 1.21.

Problema 1.9: suponiendo que la probabilidad de obtener un preparado químico por un determinado procedimiento sea de 0'01, ¿cual será el número medio de éxitos y la probabilidad de tener al menos un éxito en 200 pruebas? Obtener el valor exacto por la distribución binomial y el valor aproximado por la distribución de Poisson y compararlos.

Actividad 1.22. Si tenemos una variable aleatoria que varía de forma continua en R, habremos de definir un conjunto de intervalos de la misma para determinar las frecuencias o probabilidades de los valores en cada intervalo,



como hicimos en el Problema 1.2 con la longitud de la mano. Pero podemos definir también una *distribución de densidad probabilística* mediante una función $p:R \to R^+ + \{0\}$ que cumpla

$\int p(x)\, dx = 1$.

En este caso, la probabilidad de un intervalo [a,b[vendrá dada por

$p([a,b]) = \int_a^b p(x)\, dx$

Naturalmente, si hacemos una partición de R en un conjunto de intervalos disjuntos, la suma de sus probabilidades valdrá 1. A partir del

Teorema -1.6: para toda función integrable f y todo intervalo [a,b[de R, existe $\xi \in [a,b[$ tal que

$\int_a^b x \cdot p(x)\, dx = \xi \cdot \int_a^b p(x)\, dx$

demostrar el

Teorema 1.23: si tenemos una distribución p de densidad probabilísitca, partimos R en intervalos disjuntos $[z-\varepsilon, z+\varepsilon[$ tomando z como valor del intervalo, y definimos la media de la distribución de densidad probabilística como el límite de la media de la correspondiente distribución probabilística cuando ε tienda a cero será $\mu(X) = \int x \cdot p(x)$ dx .

Teorema 1.24: definiendo la varianza de una distribución p de densidad probabilística como $\mu((X-\mu(X))^2)$, será

$\sigma^2(X) = \int x^2 \cdot p(x)\, dx - \mu(X)^2$.

Actividad 1.23. Definimos la distribución *normal* $N(\alpha,\beta)$ por $P^N(x) = e^{-(x-\alpha)2/(2\beta 2)}/(\beta(2\pi)^{1/2})$ para todo $x \in R$.

Teniendo en cuenta el

Teorema -1.7: $\int_{-\infty}^{\infty} e^{-u^2}du = \sqrt{\pi}$, $\int_{-\infty}^{\infty} ue^{-u^2}du = 0$, $\int_{-\infty}^{\infty} u^2e^{-u^2}du = (\sqrt{\pi})/2$.
demostrar los
Teorema 1.25: $\int_{-\infty}^{\infty} P^N(x)\,dx = 1$ (y por lo tanto se trata de una distribución de densidad probabilística)
Teorema 1.26: $\mu(N(\alpha,\beta)) = \alpha$
Teorema 1.27: $\sigma(N(\alpha,\beta)) = \beta$
Escribiremos por lo tanto $N(\mu,\sigma)$ y $P^N(x) = e^{-(x-\mu)^2/(2\sigma^2)}/(\sigma(2\pi)^{1/2})$.

Actividad 1.24. Definimos la distribución normal tipificada como N(0,1), de modo que $P^N(y) = e^{-y^2/2}/(2\pi)^{1/2}$.
Trabajaremos con la tabla de la distribución normal tipificada. A partir de ésta podemos obtener fácilmente los valores de otra distribución normal mediante la normalización de su variable x, de forma que $y=(x-\mu)/\sigma$, y teniendo en cuenta que $P^N(y)=\sigma\cdot P^N(x)$.
Problema 1.10. Utilizando la tabla de la distribución normal tipificada, obtener la densidad probabilística de una distribución normal con $\mu=5$, $\sigma=2$ para x=7'4.

Actividad 1.25. La importancia de la distribución normal para el estudio de la Estadística resulta justificada por el siguiente
Teorema 1.28: si tenemos una sucesión de variables aleatorias independientes X^i con la misma media y desviación típica, $\mu(X^i)=\mu$, $\sigma(X^i)=\sigma$, y definimos $Z^n = \sum_{i=1}^{n} X^i/n$, entonces la distribución estadística de $\lim_{n\to\infty} N(Z^n)$ es la distribución normal tipificada N(0,1) (*Teorema central del límite*).
De acuerdo con este teorema, la distribución normal dará una buena aproximación de la media de un gran número de variables aleatorias equivalentes independientes, y podremos utilizarla cuando trabajemos con grandes cantidades de datos. En particular, se cumple el
Teorema 1.29: $\lim_{n\to\infty} P^{B(p,n)}(m) / P^{N(np,\sqrt{(np(1-p))})}(m) = 1$ (teorema de De Moivre; para cada valor de m, la sucesión de valores de n será n=m, m+1, m+2...)
La distribución normal es una buena aproximación a la binomial si $n\cdot p>5$ y $n\cdot q>5$.
Problema 1.11: comparar las distribuciones normal, binomial y de Poisson en los siguientes casos:
a) Aplicar la distribución normal para intentar aproximar la solución del Problema 1.9. ¿Da una buena aproximación?
b) Suponiendo que la probabilidad de obtener un preparado químico por un determinado procedimiento sea de 0'5, ¿cual será la probabilidad de tener únicamente un fracaso en 10 pruebas? Obtener el valor exacto por la distribución binomial e intentar aproximarlo por las distribuciones normal y de Poisson. ¿Cuál da una mejor aproximación?

Trabajo 1 (para su realización en equipo)**:**
Estudiar las condiciones de aproximación de las distribuciones normal y de Poisson a la distribución binomial, utilizando diferentes fuentes bibiogràfiques (por ejemplo http://www.gestiopolis.com/recursos/experto/catsexp/pagans/eco/44/distripoisson.htm

http://www.suagm.edu/paginas/japaricio/384/clase11.pdf
http://www.jstor.org/sici?sici=0003-4851(196009)31%3A3%3C737%3ATPATTP%3E2.0.CO%3B2-Q&cookieSet=1
http://leonsotelo.blogspot.com/2007/06/aproximacion-normal-binomial-poiss0n.html
http://www.mitecnologico.com/Main/AproximacionDeBinomialPorDePoisson)
Estudiar y comparar en particular el caso n=30, p=1/6. Obtener una tabla de las tres distribuciones (para valores enteros no negativos) y representarlas gráficamente en la misma figura.