DISTRIBUCION DE BERNOULLI

Una prueba de Bernoulli es aquel experimento que tiene como resultado la realización (éxito) o no (fracaso) de determinado suceso A. El espacio muestral consta de dos puntos, $w_1 = A$, $w_2 = A^c$. Podemos definir una variable aleatoria, de la forma

$$X(w_1) = 1, \quad X(w_2) = 0,$$

si $P(A) = p$ y $P(A^c) = 1-p = q$, entonces esta variable aleatoria, que se conoce como v.a. de Bernoulli, tiene por función de cuantía,

$$f_X(1) = p, \quad f_X(0) = q, \quad f_X(x) = 0, \ \forall x \in R - \{0,1\}.$$

Es muy fácil comprobar que se trata de una función de cuantía.

DISTRIBUCION BINOMIAL

Supongamos n pruebas de Bernoulli independientes, de manera que en cada una de ellas el resultado es éxito o fracaso y p y $q = 1-p$ son, respectivamente, las probabilidades de éxito o fracaso en cada una de las pruebas. El espacio muestral de las n pruebas, Ω, estará constituido por todas las posibles combinaciones de éxitos y fracasos que pueden ocurrir en las n pruebas, es decir

$$\Omega = \{w_1, \ldots, w_k\} \quad \text{con, por ejemplo,} \quad w_i = AAA^cAA^c \overset{n}{\frown} A, \quad A\text{-éxito}, \ A^c\text{-fracaso}.$$

tendremos pues ~~[illegible]~~ que todos los puntos del espacio muestral no son igualmente probables. ~~Obsérvese que~~ Así, por ejemplo, si w_i es un punto de Ω que comporta x éxitos y $n-x$ fracasos, tendremos

$$P(w_i) = p^x \cdot q^{n-x}.$$

Definamos, sobre Ω, la siguiente variable aleatoria

$$X(w_i) = \text{nº de éxitos que } w_i \text{ comporta},$$

se trata de una variable discreta y, concretamente, con un rango finito, a saber $S_d = \{0,1,2,\ldots,n\}$. [Es costumbre también escribir $X = \{0,1,\ldots,n\}$, para designar el rango de la función]. A esta variable se la conoce como variable aleatoria Binomial. Estudiemos su función de cuantía y veremos el porqué de tal nombre.

Ya sabemos que:

$$f_X(x) = P(X = x),$$

por tanto, si $x \in R - S_d$, entonces $f_X(x) = 0$, pero para $x \in S_d$, tendremos

$$x \in S_d, \quad f_X(x) = P(X = x) = P(\{w_i, \text{tales que en } w_i \text{ hay } x \text{ éxitos}\}) = P(B)$$

sabemos que cada uno de los $w_i \in B$, tienen una probabilidad, $P(w_i) = p^x \cdot q^{n-x}$, pero, ¿cuántos hay?. Los x éxitos pueden distribuirse de $\binom{n}{x}$ formas entre los n resultados, así pues B contiene $\binom{n}{x}$ elementos, y en definitiva

$$f_X(x) = \binom{n}{x} p^x q^{n-x}, \quad x \in \{0,1,\ldots,n\}.$$

Efectivamente f_X es una función de cuantía pues

a) $f_X(x) \geq 0, \ \forall x \in R$

b) $\displaystyle\sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} f_X(x) = \sum_{x=0}^{n} (p+q)^n = 1^n = 1.$

El nombre de binomial se justifica si observamos que los valores de la función de cuantía son los elementos del desarrollo del binomio $(p+q)^n$.

Es costumbre designar abreviadamente la distribución binomial de la manera $B(n,p)$, donde n y p son llamados parámetros de la distribución. Observemos que conocidos n y p la distribución binomial queda perfectamente determinada por cuanto conocemos el nº de pruebas de Bernoulli que componen la experiencia y la probabilidad de éxito en cada una de ellas.

La v.a. Binomial es de aplicación en todas aquellas experiencias que consisten en:

a) La experiencia consta de un número finito de pruebas, independientes unas de otras.

b) En cada prueba ocurre o no un determinado suceso (éxito o fracaso)

c) La probabilidad de ocurrencia del suceso es la misma en cada prueba.

en estas condiciones la $B(n,p)$ describe adecuadamente la probabilidad de obtener x éxitos en la citada experiencia (o de obtener $n-x$ fracasos).

EJEMPLOS DE UTILIZACION DE LA BINOMIAL

I) Los dados de Weldon. Consideremos un experimento consiste en lanzar un dado 12 veces y sea "exito" la aparición de un cinco o un seis. Con un dado perfecto la probabilidad de exito viene dada por $p = 1/3$ y el número de exitos debería seguir una distribución Binomial $B(12, 1/3)$. Llevamos a cabo el experimento un total de 26306 veces y recogemos en una tabla, reproducida parcialmente a continuación, las frecuencias relativas observadas y las probabilidades teóricas. De su comparación (en su día veremos que metodología emplear para llevarla a cabo) y teniendo en cuenta el elevado número de repeticiones, puede deducirse que los datos teóricos no se ajustan adecuadamente a los experimentales. Puede suponerse, a la vista de este resultado, que el dado utilizado presenta un pequeño sesgo a favor de la aparición del cinco o el seis, de tal manera que $p = 0.3377$ sería una más razonable probabilidad de exito.

DATOS DEL DADO DE WELDON

k	B(12, 1/3)	Frecuencia observada	B(12, 0.3377)
⋮	⋮	⋮	⋮
3	0.211952	0.208127	0.207736
4	0.238446	0.232418	0.238324
5	0.190757	0.197445	0.194429
6	0.111275	0.115589	0.115660
⋮	⋮	⋮	⋮

II) Supongamos que el número de "exitos" en determinada experiencia sigue una ley binomial $B(n, 0.01)$, ¿cuantas repeticiones debemos llevar a cabo para asegurar que la probabilidad de al menos un exito sea $\frac{1}{2}$ o más?

$$P(X \geq 1) = 1 - P(X=0) = 1 - (0.99)^n \geq \frac{1}{2}$$

y de aquí $-n \log(0.99) \geq \log 2$ lo que da $n \geq 70$.

III) Un problema de suministro eléctrico. Supongamos que $n=10$ trabajadores usan una fuente de energía eléctrica intermitentemente y estamos interesados en estimar la carga total esperada. Como una aproximación al problema imaginemos que en un momento determinado cada trabajador tiene la misma probabilidad p de necesitar utilizar una unidad de potencia. Si trabajan independientemente, la probabilidad de que exactamente k de ellos necesiten energía al mismo tiempo, vendrá dada por $B(n, p)$. Si por término medio un trabajador utiliza

la fuente de energía 12 minutos por hora, $p = 1/5$. La probabilidad de que siete o más trabajadores requieran corriente en el mismo momento vendrá dada por:

$$P(X \geq 7) = \sum_{x=7}^{10} \binom{10}{x} \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{10-x} = 0.0008643584 .$$

En otras palabras si la fuente de energía se ajusta a 6 unidades de potencia, la probabilidad de sobrecarga de uso será 0.00086.... lo que en términos de tiempo, tomando el minuto como unidad básica, que 1 de cada 1157 minutos podría ocurrir dicha sobrecarga, es decir 1 minuto cada 20 horas de uso. La probabilidad para 8 o más trabajadores utilizando la corriente al mismo tiempo es 0.0000779264, es decir, unas 11 veces menor.

IV) Contraste de sueros o vacunas. Supongamos que la incidencia normal de determinada enfermedad en el ganado vacuno es 0.25. Para contrastar un suero descubierto recientemente se inyectan con él n animales sanos, ¿cómo evaluar los resultados del experimento?. Para un suero sin efectividad ninguna la probabilidad de que exactamente k de los n animales inyectados permanezcan sin infección, viene dada por $B(n, 0.75)$. Para $k=n=10$ la probabilidad buscada es 0.056 y para $k=n=12$ es 0.032. Así, si entre 10 o 12 animales inyectados ninguno de ellos enferma puede interpretarse este hecho como una prueba de la efectividad del suero utilizado, aunque no se trate de una prueba concluyente. No obstante para $n=17$, la $P(X \leq 1) = 0.0501$, siendo X el número de infectados, ello puede interpretarse en el sentido de que al inyectar 17 animales si sólo uno enferma la evidencia de la efectividad del suero es mayor que cuando inyectamos a 10 y ninguno enfermó. En la misma línea de razonamiento, para $n=23$, $P(X \leq 2) = 0.0492$ y por tanto tendremos una evidencia más fuerte de la bondad del suero cuando al inyectar a 23 individuos sólo 2 de ellos enfermen.

LA DISTRIBUCIÓN DE POISSON (Como una aproximación de la binomial)

En la fabricación de botellas de vidrio, se consideran defectuosas aquellas en las que aparecen partículas sólidas, "piedras", que estaban previamente diluidas en la masa de vidrio fundido. Supongamos que tenemos 100 kgs. de masa, que cada botella requiere 1 kgs para su fabricación y que hay x "piedras" en los 100 kgs. de masa. ¿Cual será el porcentaje de botellas defectuosas?. La respuesta x % ignora la posibilidad de que en una botella aparezca más de una "piedra". Para contestar a la pregunta podemos recurrir a un <u>modelo probabilístico a base de urnas</u>.

Bajo el supuesto de que cada "piedra" tiene la misma probabilidad de aparecer en cada botella y esto independientemente de lo que ocurra con las otras "piedras", podemos reducir el problema a lo siguiente: se colocan aleatoriamente n bolas en N urnas, ¿cual es la probabilidad de que una urna elegida al azar contenga exactamente k bolas?. Como ya sabemos

casos posibles N^n
casos favorables $(N-1)^{n-k}$

esto para k bolas fijas, pero tengamos en cuenta que estas k bolas pueden elegirse de $\binom{n}{k}$ formas distintas, en consecuencia la probabilidad buscada es

$$W_k = \binom{n}{k}\frac{(N-1)^{n-k}}{N^n} = \binom{n}{k}\frac{1}{N^k}\cdot\left(1-\frac{1}{N}\right)^{n-k}$$

Si la producción de las N botellas requiere M quintales de masa de vidrio, tendremos, en este caso, $N=100M$ y $n=xM$. Como estamos interesados en el porcentaje de botellas defectuosas en un largo periodo de fabricación, supondremos que M es muy grande. Sea $\lambda = \frac{x}{100}$, entonces

$$\left.\begin{array}{l} N=100M \\ n=xM \end{array}\right\} N = \frac{100}{x}\cdot n = \frac{n}{\lambda}$$

, sustituyendo en W_k, tendremos

$$W_k = \frac{1}{k!}\cdot\prod_{j=0}^{k-1}(n-j)\cdot\left(\frac{\lambda}{n}\right)^k\cdot\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!}\cdot\left(1-\frac{\lambda}{n}\right)^{n-k}\cdot\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right).$$

que M sea grande, supone que n es grande y sabemos que

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

y que

$$\lim_{n\to\infty}\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right) = 1$$

entonces

$$\lim_{n\to\infty} W_k = \frac{\lambda^k}{k!}e^{-\lambda} \qquad k=0,1,\ldots$$

A la distribución

$$P_k = P(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \qquad k=0,1,\ldots$$

se la conoce con el nombre de distribución de Poisson. Se trata de una probabilidad por cuanto

$$\sum_{k\geq 0} P_k = \sum_{k\geq 0}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\cdot\sum_{k\geq 0}\frac{\lambda^k}{k!} = e^{-\lambda}\cdot e^{\lambda} = 1.$$

Volvamos de nuevo al caso práctico inicial. La relación entre frecuencia relativa y probabilidad hace que la proporción de botellas defectuosas frente al total fabricado sea aproximadamente igual a la probabilidad de que haya al menos una botella defectuosa y esta probabilidad puede aproximarse mediante $1-W_0 \simeq 1-e^{-\lambda}$ y como $\lambda = \frac{x}{100}$, el porcentaje de piezas defectuosas será $100\left(1-\exp\left(-\frac{x}{100}\right)\right)$. Si x es muy pequeño, esta cantidad es prácticamente igual a x, pero para valores grandes x no ocurre así. Por ejemplo, para $x=100$, el porcentaje de botellas defectuosas no es 100 sino $63.21\% = 100(1-e^{-1})$, cantidad que es de esperar que no ocurra evitando que x sea tan grande. Si por ejemplo $x=30$, el porcentaje será 25.92%

<u>OBSERVACIÓN</u> Puede asimilarse a una binomial razonando así. La probabilidad de que una bola sea introducida en una urna es $\frac{1}{N}$, y la contraria $\left(1-\frac{1}{N}\right)$; por tanto la de obtener k-éxitos (éxito = bola en la urna en cuestión) será

$$\binom{n}{k}\left(\frac{1}{N}\right)^k\left(1-\frac{1}{N}\right)^{n-k}$$

en lugar de 30%, que obtendríamos razonando como al principio. Evidentemente si el número de "piedras" en la masa es grande es mucho más económico, si el proceso lo permite, fabricar botellas pequeñas, sino hay forma posible de limpiar la masa. Así por ejemplo, fabricando botellas que requieran sólo 0.25 kgs de masa en lugar de 1 Kg., el nuevo valor de λ es $\lambda' = x/400$ y la proporción de defectuosas, para $x=30$, baja del 25.92% al 7.23%. La teoría de la probabilidad proporciona, como hemos visto, sugerencias interesantes para problemas prácticos de producción. (**) ver APENDICE

LA DISTRIBUCIÓN DE POISSON (Como solución a un problema físico-químico)

Supongamos una cierta masa de sustancia radiactiva cuyo periodo de vida media es elevado. Supongamos además que las siguientes hipótesis son ciertas

1.- Si $t_1<t_2<t_3$ y $A_k(t_1,t_2)$ designa el suceso "durante el intervalo de tiempo (t_1,t_2) ocurren k desintegraciones", entonces los sucesos $A_k(t_1,t_2)$ y $A_\ell(t_2,t_3)$ son independientes para cualesquiera valores no negativos de k y l.

2.- Los sucesos $A_k(t_1,t_2)$, $k=0,1,\ldots$ constituyen una partición del espacio muestral. Para k dado, $P(A_k(t_1,t_2))$ depende sólo de la diferencia t_2-t_1. En otras palabras, el proceso de desintegración radiactiva es homogéneo con respecto al tiempo. $W_k(t)$ designará la probabilidad de k desintegraciones durante un intervalo de tiempo de longitud t ($t_2-t_1=t$).

3.- Si t es suficientemente pequeño, la probabilidad de que durante un intervalo de tiempo t ocurra más de una desintegración es despreciablemente pequeña comparada con la probabilidad de que ocurra exactamente una. Es decir

$$\lim_{t\to 0} \frac{1-W_0(t)-W_1(t)}{W_1(t)} = 0, \qquad (1)$$

o lo que es lo mismo

$$\lim_{t\to 0} \frac{1-W_0(t)}{W_1(t)} = 1 \qquad (2)$$

En otras palabras: la probabilidad de que ocurra al menos una desintegración es, en el límite, igual a la probabilidad de que ocurra exactamente 1.

Claramente $W_0(0)=1$ y $W_k(0)=0$, $k\geq 1$. Además $W_0(t)$ es una función monótona decreciente. De aquí y de las condiciones 1 y 2 tenemos

$$W_0(t+s) = P(A_0(t_1,t_3)) = P\big(A_0(t_1,t_2)\cap A_0(t_2,t_3)\big) = P(A_0(t_1,t_2))\cdot P(A_0(t_2,t_3)) =$$
$$= W_0(t)\cdot W_0(s) \quad \text{para } t_2-t_1=t \text{ y } t_3-t_2=s.$$

se trata pues de una función aditiva, si exigimos la derivabilidad de la función (en 0), tendremos

$$W_0(t) = e^{-\mu t} \quad \text{con } \mu>0. \qquad (3)$$

En orden a determinar las funciones $W_k(t)$ observemos que

$$\lim_{\Delta t\to 0} \frac{W_k(\Delta t)}{\Delta t} = 0, \quad k\geq 2 \qquad (5)$$

En efecto

$$0 \leq W_k(t) \leq \sum_{k\geq 2} W_k(t) = 1-W_0(t)-W_1(t), \quad \forall k\geq 2$$

y de aquí

$$0 \leq \frac{W_k(\Delta t)}{\Delta t} \leq \frac{1-W_0(\Delta t)-W_1(\Delta t)}{\Delta t}\cdot\frac{W_1(\Delta t)}{W_1(\Delta t)} = \frac{1-W_0(\Delta t)-W_1(\Delta t)}{W_1(\Delta t)}\cdot\frac{W_1(\Delta t)}{\Delta t} \qquad (4)$$

de (2), tenemos

$$\lim_{t\to 0}\frac{(1-W_0(t))/t}{W_1(t)/t} = 1 \longrightarrow \lim_{t\to 0}\frac{1-W_0(t)}{t} = \lim_{t\to 0}\frac{W_1(t)}{t} = W_1'(0)$$

pero el primer miembro de esta última igualdad es $-W_0'(0)$ que existe y por tanto $\lim_{t\to 0}\frac{W_1(t)}{t}$ es una cantidad finita. Pasando al límite en (4) y teniendo en cuenta (1) llegamos al resultado buscado. Ahora bien, como para $k\geq 1$, $W_k(0)=0$, (5) puede escribirse

$$W_k'(0) = 0, \quad k\geq 2 \qquad (6)$$

lo que prueba la existencia de $W_k'(0)$.

Por otro lado consideremos ahora el suceso "k desintegraciones ocurren durante el intervalo de tiempo $(0,t+\Delta t)$". Este suceso puede darse de tres formas distintas:

a) $k-1$ desintegraciones entre 0 y t y una entre t y $t+\Delta t$,

b) k desintegraciones entre 0 y t y 0 entre t y $t+\Delta t$,

c) a lo sumo $k-2$ desintegraciones entre 0 y t y al menos 2 entre t y $t+\Delta t$.

por tanto, de acuerdo con las hipótesis 1 y 2, tendremos

$$W_k(t+\Delta t) = W_{k-1}(t)\cdot W_1(\Delta t) + W_k(t)\cdot W_0(\Delta t) + R \qquad (7)$$

donde $R = o(\Delta t)$, de acuerdo con la hipótesis 3 y la relación (5). De aquí

$$W_k(t+\Delta t) - W_k(t) = W_k(t)\,[W_0(\Delta t)-1] + W_{k-1}(t)\cdot W_1(\Delta t) + R$$

dividiendo por Δt y pasando al límite

$$\lim_{\Delta t\to 0}\frac{W_k(t+\Delta t)-W_k(t)}{\Delta t} = W_k(t)\lim_{\Delta t\to 0}\frac{W_0(\Delta t)-1}{\Delta t} + W_{k-1}(t)\lim_{\Delta t\to 0}\frac{W_1(\Delta t)}{\Delta t} + \overbrace{\lim_{\Delta t\to 0}\frac{R}{\Delta t}}^{0 \text{ por } (5)}$$

teniendo en cuenta (3) y el hecho de que $W_0'(0) = -W_1'(0)$, queda

$$W_k'(t) = \mu\left(W_{k-1}(t) - W_k(t)\right) \quad , \; k = 1,2,\ldots \qquad (8)$$

para resolver esta ecuación diferencial hagamos

$$V_k(t) = W_k(t) \cdot e^{\mu t}$$

entonces derivando

$$V_k'(t) = W_k'(t) \cdot e^{\mu t} + W_k(t) \cdot \mu \cdot e^{\mu t}$$

y de aquí

$$W_k'(t) = \frac{V_k'(t)}{e^{\mu t}} - W_k(t) \cdot \mu$$

entonces, de (8)

$$\frac{V_k'(t)}{e^{\mu t}} = \mu \cdot W_{k-1}(t) \rightarrow V_k'(t) = \mu \cdot W_{k-1}(t) \cdot e^{\mu t} = \mu \cdot V_{k-1}(t) \quad , \; k = 1,2,\ldots$$

De $W_0(t) = e^{-\mu t}$ se sigue $V_0(t) = 1$ y de aquí

$$V_1(t) = \mu t \quad , \quad V_2(t) = \frac{\mu^2 t^2}{2}$$

y en general

$$V_k(t) = \frac{(\mu t)^k}{k!}.$$

Por tanto

$$W_k(t) = \frac{(\mu t)^k}{k!} \cdot e^{-\mu t} \quad , \; k = 0,1,\ldots$$

Es decir, que el número de desintegraciones durante un intervalo de tiempo t, dadas las hipótesis 1-3, tiene una distribución de Poisson con parámetro μt.

OBSERVACIÓN

La distribución de Poisson que acabamos de obtener como solución al problema específico de las desintegraciones de sustancias radiactivas es aplicable, como modelo probabilístico, a un sinfín de situaciones a condición de que se verifiquen, debidamente modificadas, las hipótesis 1-3 citadas al comienzo del párrafo. Así por ejemplo, si en lugar de intervalos de longitud t tenemos dominios de área S, o volumen V y estamos interesados en calcular la

probabilidad de encontrar k partículas en un dominio determinado, podemos recurrir al modelo de Poisson a condición de que:

1) El proceso estudiado sea homogéneo. Es decir, la probabilidad de encontrar los k puntos en un dominio determinado depende sólo del área o volumen del dominio y no de su forma.

2) Si el dominio es pequeño (en cuanto a área o volumen), la probabilidad de encontrar más de un punto en dicho dominio es pequeña comparada con el tamaño del recinto o incluso si la comparamos con la probabilidad de encontrar una sola partícula.

3) Dominios disjuntos son mutuamente independientes.

NOTA HISTÓRICA. OÍDO!!

Simeon D. Poisson vivió de 1781 a 1840. Su distribución de probabilidad la presentó por primera vez en 1837 en un libro titulado "Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités". Toma ya fardel!

ALGUNOS EJEMPLOS DE OBSERVACIONES QUE SE AJUSTAN A UNA DISTRIBUCIÓN DE POISSON.

I) Desintegraciones radioactivas. Ya sabemos en este caso del ajuste del fenómeno a la distribución de Poisson, pero estudiemos con más detalle un experimento famoso. En 1920, Rutherford, Chadwick y Ellis estudiaron, en Cambridge una sustancia radioactiva durante $N = 2608$ intervalos de 7.5 segundos cada uno, obteniendo el número de partículas que alcanzan un contador Geiger. En la tabla se recogen el número N_k de periodos con exactamente k partículas.

k	N_k	$N.P(X=k)$	k	N_k	$N.P(X=k)$
0	57	54.399	5	408	393.515
1	203	210.523	6	273	253.817
2	383	407.361	7	139	140.325
3	525	525.496	8	45	67.882
4	532	508.418	9	27	29.189
			$k \geq 10$	16	17.075
			Total	2608	2608.000

El número total de partículas fue de $T = \sum k N_k = 10094$ y por término medio en cada intervalo $T/N = 3.870$. Se obtuvieron entonces los valores teóricos $N.P(X=k)$, donde $P(X=k)$ se obtuvo a partir de una distribución de Poisson con parámetro $\lambda = 3.870$, es decir partículas por unidad de tiempo empleada, a saber intervalos de 7.5 segundos. La comparación de los valores teóricos con los experimentales, con métodos que no hacen al caso ahora, sugieren que el experimento en cuestión puede explicarse correctamente mediante la correspondiente distribución de Poisson. (7)

II) <u>Bombas volantes sobre Londres.</u> Como ejemplo de una distribución espacial de puntos aleatorios consideremos las estadísticas de bombas volantes que cayeron sobre Londres durante la Segunda Guerra Mundial. Se divide la ciudad en $N = 576$ áreas pequeñas de superficie $S = \frac{1}{4}\ km^2$ cada una. La tabla presenta el número N_k de áreas con exactamente k bombas. El número total de impactos fue de $T = \sum k \cdot N_k = 537$ y la media de impactos por área $T/N = 0.9323$ que tomaremos

k	0	1	2	3	4	5 y más
N_k	229	211	93	35	7	1
$N.P(X=k)$	226.74	211.39	98.54	30.62	7.14	1.57

como valor de λ para la distribución de Poisson correspondiente. La concordancia entre los datos teóricos y los experimentales es más que aceptable.

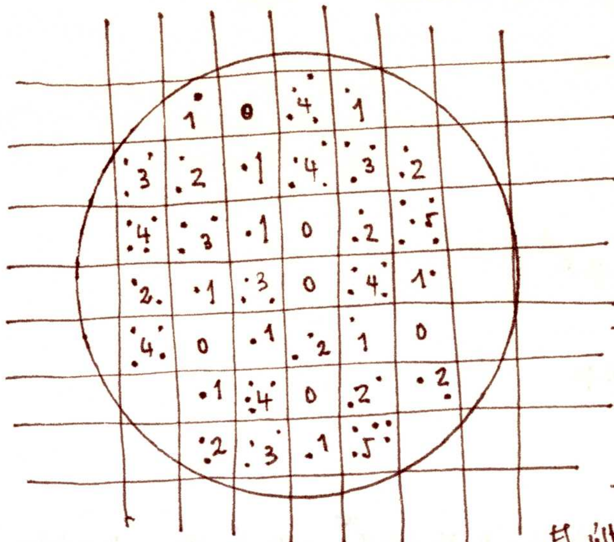
Es interesante señalar que existe una creencia general en el sentido de que los impactos tendían a agruparse. Si esto hubiera sido cierto hubieran existido gran cantidad de áreas con muchos o muy pocos impactos y por contra gran deficiencia de áreas con un número intermedio de impactos. La tabla presentada indica la homogeneidad y aleatoriedad de las áreas y deshace la creencia general que tiende a identificar la aleatoriedad como una cierta tendencia al agrupamiento.

III) <u>Conexiones telefónicas incorrectas.</u> La tabla presenta estadísticas acerca de las conexiones telefónicas a un número equivocado. Un total de $N = 267$ números fueron observados; N_k indica cuantos

k	N_k	$N.P(X=k)$	k	N_k	$N.P(X=k)$
0-2	1	2.05	11	20	24.34
3	5	4.76	12	18	17.72
4	11	10.39	13	12	11.92
5	14	18.16	14	7	7.44
6	22	26.45	15	6	4.33
7	43	33.03	≥ 16	2	4.65
8	31	36.09	Total	267	267.00
9	40	35.04			
10	35	30.63			

números tuvieron exactamente k conexiones erróneas. El número medio de estas conexiones por número de teléfono considerado fue de 8.74 que se toma como valor de λ para la distribución de Poisson que nos permite obtener $P(X=k)$.

IV) <u>Recuentos de hematíes y bacterias.</u> (8)



La figura reproduce una placa de Petri con colonias de bacterias, visibles al microscopio como puntos oscuros. La placa se divide en cuadrados pequeños, la tabla reproduce el número de cuadrados observados con exactamente k puntos, en varios experimentos llevados a cabo.

k	0	1	2	3	4	5	6
N_k	5	19	26	26	21	13	8
Poisson	6.1	18.0	26.7	26.4	19.6	11.7	9.5
N_k	8	16	18	15	9	7	
Poisson	6.8	16.2	19.2	15.1	9.0	6.7	
N_k	60	80	45	16	9		
Poisson	62.6	75.8	45.8	18.5	7.3		

El último valor incluye todos aquellos cuadrados con k o más colonias.

LA DISTRIBUCIÓN BINOMIAL NEGATIVA

Consideremos n pruebas de Bernoulli, es decir, pruebas cuyo resultado es éxito o fracaso con probabilidades p y $q=1-p$, respectivamente. Supongamos además que se trata de pruebas independientes. Estamos interesados en conocer cuando ocurrirá el r-ésimo éxito. Aquí r es un entero positivo fijo. Con las condiciones planteadas, es obvio que la probabilidad de que el r-ésimo éxito ocurra en la prueba $\nu \leq n$ es independiente de n y depende tan sólo de ν, r y p. Puesto que necesariamente $\nu \geq r$, podemos escribirlo de la forma $\nu = r+k$. La probabilidad de que el r-ésimo éxito ocurra en la prueba $r+k$ $(k=0,1,\ldots)$ la denotaremos mediante $f(k;r,p)$. Esta probabilidad es igual a la probabilidad de que ocurran exactamente k fracasos antes del r-ésimo éxito. Este suceso ocurre si y sólo si, entre las $r+k-1$ pruebas hay exactamente k fracasos y en la siguiente, a saber la $(r+k)$-ésima, el resultado de la prueba es un éxito, la probabilidad de un suceso tal como el descrito viene dada por

$\binom{r+k-1}{k} p^{r-1}\cdot q^k$ para los k fracasos en las $r+k-1$ pruebas, y

p para el último éxito

en consecuencia

$$f(k;r,p) = \binom{r+k-1}{k} p^r \cdot q^k \qquad k=0,1,2,\ldots$$

Que suele también escribirse de la forma

$$f(k;r,p) = \binom{-r}{k} p^r \cdot (-q)^k \qquad k=0,1,2,\ldots$$

de donde le viene el nombre de Binomial negativa. Téngase en cuenta que:

$$\binom{-r}{k} = \frac{(-r)(-r-1)\cdots(-r-k+1)}{k!} = \frac{(-r)(-(r+1))\cdots(-(r+k-1))}{k!} = \frac{(-1)^k r.(r+1)\cdots(r+k-1)}{k!} =$$

$$= \frac{(-1)^k (r-1)!\, r.(r+1)\cdots(r+k-1)}{(r-1)!\; k!} = (-1)^k \frac{(r+k-1)!}{(r-1)!\, k!} = (-1)^k \binom{r+k-1}{k}.$$

Supongamos ahora que la sucesión de pruebas de Bernoulli la continuamos el término de veces necesarias para conseguir los r éxitos. En teoría puede ocurrir que el éxito r-ésimo no aparezca nunca y debamos prolongar nuestra sucesión hasta infinito.

Tenemos así descrita una variable aleatoria discreta X, que toma los valores $\{0,1,2,\ldots\}$ y que representa el número de pruebas adicionales a r que debemos realizar para alcanzar

r éxitos si la probabilidad de cada éxito es p, $q=1-p$ la de obtener un fracaso, y las pruebas se llevan a cabo de forma independiente. La probabilidad de que $X=k$, viene dada por

$$P(X=k) = f(k;r,p) = \binom{r+k-1}{k} p^r\cdot q^k = \binom{-r}{k} p^r (-q)^k.$$

No debe existir duda alguna respecto al hecho de que la función $f(k;r,p)$ es una función de cuantía, por cuanto:

1) $f(k;r,p) \geq 0 \quad ,\forall k$

2) $f(k;r,p)$ es discreta por cuanto toma valores distintos de cero para $k=0,1,2,\ldots$ que es un conjunto discreto de valores

3) $f(k;r,p)$ es normada. En efecto

$$\sum_{k\geq 0} f(k;r,p) = 1$$

ya que si recordamos el desarrollo en serie de Mc Laurin de la función $(1-x)^{-n}$ tenemos:

$$\frac{1}{(1-x)^n} = \sum_{j\geq 0} \binom{n+j-1}{j} x^j \qquad |x|<1$$

en nuestro caso

$$\sum_{k\geq 0} \binom{r+k-1}{k} p^r q^k = p^r \sum_{k\geq 0} \binom{r+k-1}{k} q^k = p^r \cdot \frac{1}{(1-q)^r} = 1 \qquad |q|<1.$$

La distribución de probabilidad Binomial Negativa es conocida también como distribución de *Pascal*. Para $r=1$ la distribución se conoce con el nombre de *distribución geométrica* y su función de cuantía se reduce a $f(k;1,p) = p\cdot q^k$, $k=0,1,\ldots$

EJEMPLOS

I) *El problema de las cajas de cerillas de Banach*. Un conocido matemático lleva siempre una caja de cerillas en cada uno de los dos bolsillos de su chaqueta. Cuando necesita una cerilla selecciona al azar uno de los bolsillos de manera que las sucesivas elecciones constituyen pruebas de Bernoulli con $p=\frac{1}{2}$. Supongamos que inicialmente cada caja contiene exactamente N cerillas y consideremos el momento en el cual, por primera vez, el matemático descubre que una de las cajas stá vacía. En ste momento la otra caja puede contener $0, 1, 2, \ldots, N$ cerillas, y vamos a designar mediante u_r la correspondiente probabilidad. Identificaremos el "éxito" con la elección del bolsillo izquierdo. Este bolsillo stará vacío, o mejor, descubriremos que stá vacío justo en el momento que el bolsillo derecho contiene exactamente r cerillas, si y sólo si exactamente $(N-r)$ fracasos han precedido al $(N+1)$-simo éxito. La probabilidad de ste suceso s $f(N-r; N+1, \frac{1}{2})$. El mismo argumento puede aplicarse respecto del bolsillo derecho y por tanto la probabilidad buscada s

$$u_r = 2f\left(N-r; N+1, \tfrac{1}{2}\right) = 2 \cdot \binom{2N-r}{N-r}\left(\frac{1}{2}\right)^{N+1} \cdot \left(\frac{1}{2}\right)^{N-r} = \binom{2N-r}{N} 2^{-2N+r}.$$

Por ejemplo, para $N=50$ y $r=4$, $u_r = 0.074790$ y para $r=29$, $u_r = 0.000232$.

II) <u>Generalización: Tenis de mesa</u>. La naturaleza del problema anterior queda más de manifiesto cuando se atribuyen diferentes probabilidades a cada una de las cajas. Interpretaremos sto de forma diferente. Supongamos que dos amigos, Pedro y Pablo, juegan a un juego que puede ser tratado como una sucesión de pruebas de Bernoulli en las que las probabilidades p y q sirven como medida de la destreza de los jugadores. En el tenis de mesa el jugador que primero llega a 21 gana la partida. Por comparación con el ejemplo precedente consideremos la situación general en la que $2\nu+1$ éxitos individuales se requieren. El juego finaliza, en el mejor de los casos en la prueba $2\nu+1$ (caso de 21 a 0) o bien, en la peor de las situaciones, en la prueba $4\nu+1$ (20 a 21). Designemos por a_r la probabilidad de que Pedro gane en la prueba (jugada) $4\nu+1-r$. Esto ocurrirá si y sólo si, en las $4\nu-r$ primeras jugadas Pedro ha obtenido 2ν tantos (éxitos) y además gana la jugada siguiente. Así pues

$$a_r = f(2\nu-r; 2\nu+1, p) = \binom{4\nu-r}{2\nu} p^{2\nu+1} q^{2\nu-r}. \qquad (1)$$

La probabilidad de que Pedro gane viene dada por $a_0 + a_1 + \cdots + a_{2\nu}$. La probabilidad de que el juego termine (gana uno u otro) en la jugada $4\nu+1-r$ viene dada por $a_r + b_r$, donde b_r, probabilidad de que gane Pablo en la jugada $4\nu+1-r$, viene dada por (1) con p y q intercambiados.

Para $2\nu = N$, $p = q = \frac{1}{2}$, las probabilidades $a_r + b_r$ se reducen a las u_r del ejemplo anterior.

<u>Nota</u>.- El ejemplo de las cerillas de Banach no se debe a Banach en contra de lo que pudiera ~~parecer~~ parecer. Tiene su ~~origen~~ origen en una referencia humorística al hábito de fumador de Banach hecha por H. Steinhaus en un acto celebrado en honor de Banach. La anécdota se hizo famosa y de aquí que haya conservado el nombre hasta ahora.

<u>LA DISTRIBUCIÓN HIPERGEOMÉTRICA</u>

Supongamos n bolas de una urna de las cuales n_1 son rojas y $n_2 = n - n_1$ son negras. Se extraen r al azar. Buscamos la probabilidad de que en las r bolas extraídas haya exactamente k rojas y designaremos por q_k dicha probabilidad. k s cualquier valor entre 0 y n_1 o r si ste s más pequeño.

Para encontrar q_k, observemos que el grupo elegido contiene k bolas rojas y $r-k$ negras. Las rojas pueden elegirse de $\binom{n_1}{k}$ formas diferentes, mientras que las negras pueden serlo de $\binom{n-n_1}{r-k}$ formas distintas. Puesto que la elección de las k rojas puede combinarse con cualquiera de las posibles elecciones de las $r-k$ negras, tendremos

$$q_k = \frac{\binom{n_1}{k}\binom{n-n_1}{r-k}}{\binom{n}{r}}$$

Si con las condiciones enunciadas definimos una variable X cuyos valores representan el número de bolas rojas en la muestra extraída, sta variable será una variable aleatoria discreta cuyas probabilidades, $P(X=k)$ vienen dadas por

$$P(X=k) = q_k = \frac{\binom{n_1}{k}\binom{n-n_1}{r-k}}{\binom{n}{r}} \qquad \begin{array}{ll} k = 0, 1, \ldots, r & \text{si } r \leq n_1. \\ k = 0, 1, \ldots, n_1 & \text{si } n_1 < r \leq n. \end{array}$$

Obsérvese que

$$\sum_{k=0}^{r} P(X=k) = \sum_{k=0}^{r} q_k = \frac{1}{\binom{n}{r}} \sum_{k=0}^{r} \binom{n_1}{k}\binom{n-n_1}{r-k} = \frac{1}{\binom{n}{r}} \cdot \binom{n}{r} = 1.$$

<u>EJEMPLOS</u>.

I) El senado de los EE.UU. stá compuesto por 100 senadores, 2 por cada uno de los 50 stados. Se elige un grupo de 50 senadores y se quiere conocer la probabilidad de que un estado determinado no ste representado en el grupo y la de que en el grupo haya 1 senador de ste estado o los 2 senadores del stado en cuestión.

Aquí asimilamos senador del stado en cuestión a "rojo", con lo que $n_1 = 2$, $r = 50$ y $n = 10$ por tanto

$$q_0 = q_2 = \frac{50 \cdot 49}{100 \cdot 99} = 0.24747\ldots \quad y \quad q_1 = \frac{50}{99} = 0.50505\ldots$$

II) <u>Estimación del tamaño de una población animal a partir de datos de recaptura</u>. Supongamos que 1000 peces capturados en un lago son marcados mediante una mancha roja y se les arroja nuevamente al lago. Después de cierto tiempo se lleva a cabo una nueva captura de otros 1000 peces y 100 de entre ellos llevan la mancha roja. ¿Qué conclusiones podemos extraer respecto al número de peces en el lago? Este es un problema típico de <u>estimación estadística</u>. No es momento de profundizar y describir los distintos métodos que podrían utilizarse para su resolución, nos limitaremos a demostrar cómo la distribución hipergeométrica nos proporciona una clave para la solución de nuestro problema. Supondremos naturalmente que las dos capturas pueden ser consideradas como muestras aleatorias de la población total de peces del lago. (En la práctica semejante suposición excluye situaciones en las que las dos capturas se hacen en una misma localidad y en un corto periodo de tiempo). Supondremos también que el número de peces en el lago no cambia entre las dos capturas.

Generalizemos el problema admitiendo tamaños arbitrarios de las muestras. Sean

n = el número total de peces en el lago (desconocido)

n_1 = número de peces en la primera captura. Harán el papel de bolas rojas.

r = número de peces en la segunda captura.

k = número de peces con mancha roja en la segunda captura.

$q_k(n)$ = probabilidad de que la segunda captura contenga exactamente k peces con mancha roja.

Con esta formulación es obvio que $q_k(n)$ viene dada por la correspondiente probabilidad hipergeométrica, a saber

$$q_k(n) = \frac{\binom{n_1}{k}\binom{n-n_1}{r-k}}{\binom{n}{r}}.$$

En la práctica n_1, r y k pueden conocerse por observación, pero n es desconocida. Nótese que consideramos n como un número fijo desconocido cuyo valor en modo alguno depende del azar. Sabemos que n_1+r-k peces han sido capturados en el lago entre las dos capturas llevadas a cabo y que obviamente $n \geq n_1+r-k$. Esto es todo cuanto sabemos con certeza. En nuestro ejemplo $n_1 = r = 1000$ y $k = 100$; es teóricamente posible que $n = 1900$, pero si comenzamos con una hipótesis de este tipo, llegaremos a la conclusión de que ha ocurrido un suceso con una probabilidad extraordinariamente pequeña. En efecto, suponiendo que $n = 1900$, la probabilidad $q_k(n)$ vendrá dada, con los datos concretos del problema, por

$$q_k(1900) = \frac{\binom{1000}{100}\binom{900}{900}}{\binom{1900}{1000}} = \frac{(1000!)^2}{100!\ 1900!}$$

utilizando la fórmula de Stirling, $n! \sim \sqrt{2\pi}\, n^{n+\frac{1}{2}} \cdot e^{-n}$, para aproximar los factoriales llegaremos a que $q_k(1900) \sim 10^{-430}$ y en una situación así el sentido común nos lleva a rechazar la hipótesis como carente de todo sentido. Un razonamiento similar nos induciría a rechazar hipótesis del tipo, n es muy grande, por ejemplo $n = 10^6$. Esta consideración nos conduce a buscar un valor de n para el cual $q_k(n)$ alcance su mayor valor, puesto que en este caso nuestra observación tendría la mayor probabilidad de producirse. Para cualquier conjunto particular de observaciones n_1, r, k, el valor de n para el que $q_k(n)$ es más grande se designa mediante $\hat{n}$ y se le llama <u>estimación máximo verosímil</u> de n. Con vistas a encontrar $\hat{n}$ consideremos el cociente

$$\frac{q_k(n)}{q_k(n-1)} = \frac{(n-n_1)(n-r)}{(n-n_1-r+k)n} = \frac{n^2-n_1n-nr+n_1r}{n^2-n_1n-nr+nk}$$

Este cociente será mayor o menor que la unidad según que $nk < n_1r$ ó $nk > n_1r$. Esto significa que a medida que n aumenta la sucesión $q_k(n)$ crece primero para decrecer después alcanzando su máximo cuando n sea igual al mayor entero menor que n_1r/k. En nuestro caso particular $\hat{n} = 10.000$.

El verdadero valor de n puede ser mayor o menor que el encontrado, y estamos interesados también en dar unos límites dentro de los cuales sea razonable encontrar el verdadero valor de n. Supongamos para ello que n es más pequeño que 8.500. Sustituimos en la correspondiente expresión de $q_k(n)$ y calculamos la probabilidad de que la segunda captura contenga a lo sumo 100 peces marcados. Esta probabilidad viene dada por

$$x = q_0 + q_1 + \text{—} + q_{100}$$

y es igual a 0.04. Análogamente para $n = 12000$, la probabilidad de que la segunda captura contenga al menos 100 peces marcados es alrededor de 0.03. Con semejantes probabilidades es razonable pensar que el verdadero valor de n será alguno de los comprendidos entre 8500 y 12000.

APENDICE

Cont. de la DISTRIBUCION DE POISSON como aproximación a la binomial

Hemos resuelto una situación muy concreta, a saber, la fabricación de botellas de vidrio y la determinación aproximada del número de botellas defectuosas. Ahora bien, ¿cómo podemos utilizar esta aproximación ante una situación genérica definida mediante una binomial cualquiera $B(n,p)$?. La forma de actuar quedará explicitada en cuanto pongamos de manifiesto la relación existente entre los parámetros n,p y λ de la Binomial y su aproximación de Poisson, respectivamente. Volviendo al ejemplo, recordemos que $\lambda = \frac{x}{100}$ y que por otra parte $N = n/\lambda$, o bien

$$\lambda = \frac{n}{N} = n \cdot \left(\frac{1}{N}\right)$$

En esta Binomial $p = \frac{1}{N}$ y por tanto $\lambda = n \cdot p$, este resultado no se da por casualidad, sino que esta es la relación que existe, en general, entre el parámetro λ y los parámetros n, p de la Binomial. Destaquemos además el hecho de que n era grande y $p = \frac{1}{N}$ era pequeño, por cuanto que habíamos señalado que la aproximación era tanto mejor cuanto mayor fuera el número de botellas fabricadas, N.

Podíamos, de hecho, haber planteado el problema de forma genérica, a continuación lo haremos, pero puede resultar más ilustrativo hacerlo mediante un caso práctico.

Sea ahora $B(n,p)$ una Binomial tal que n es muy grande y p es muy pequeño, de manera que pueden planteársenos dificultades a la hora de determinar las probabilidades $P(X=k)$, $k=0,1,\ldots,n$. Hagamos $\lambda = n p$. Sabemos que

$$\left.\begin{array}{l} P(X=0) = (1-p)^n = \left(1-\frac{\lambda}{n}\right)^n \\ \text{para } n \text{ grande} \\ \lim\limits_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda} \end{array}\right\} \quad P(X=0) \simeq e^{-\lambda} \qquad (1)$$

Por otra parte

$$\frac{P(X=k)}{P(X=k-1)} = \frac{(n-k+1)\,p}{k q} = \frac{\lambda - (k-1)\,p}{k q}$$

que para p pequeño (entonces q estará muy próximo a la unidad), puede aproximarse mediante

$$\frac{P(X=k)}{P(X=k-1)} \simeq \frac{\lambda}{k} \qquad (2)$$

de (1) y (2), tenemos

$$P(X=1) \simeq \lambda \cdot P(X=0) \simeq \lambda \cdot e^{-\lambda}$$

$$P(X=2) \simeq \frac{\lambda}{2} \cdot P(X=1) \simeq \frac{\lambda^2}{2} \cdot e^{-\lambda}$$

$$P(X=3) \simeq \frac{\lambda}{3} \cdot P(X=2) \simeq \frac{\lambda^3}{2\cdot 3} \cdot e^{-\lambda}$$

$$\ldots\ldots\ldots$$

$$P(X=k) \simeq \frac{\lambda}{k} \cdot P(X=k-1) \simeq \frac{\lambda^k}{k!}\, e^{-\lambda}$$

En resumen las probabilidades binomiales $B(n.p)$ pueden aproximarse mediante las probabilidades de una distribución de Poisson, $P(\lambda)$, con $\lambda = n \cdot p$, siendo la aproximación tanto mejor cuanto mayor sea n y menor p.

$$\boxed{P(X=k) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \simeq \frac{\lambda^k}{k!}\, e^{-\lambda}} \qquad \lambda = n.p$$

EJEMPLOS de aproximación

I) Cumpleaños. ¿Cuál es la probabilidad, p_k, de que en una compañía de 500 artistas exactamente k celebren su cumpleaños el día de Año Nuevo? En el supuesto de la elección aleatoria de los 500 artistas, la probabilidad podrá obtenerse a partir de una Binomial $B\left(500, \frac{1}{365}\right)$, o bien aproximarla mediante una Poisson, $P(\lambda)$, con $\lambda = \frac{500}{365} = 1.3699\ldots$. A continuación se presentan las probabilidades correctas y sus aproximaciones.

k	0	1	2	3	4	5	6
Binomial	0.2537	0.3484	0.2388	0.1089	0.0372	0.0101	0.0023
Poisson	0.2541	0.3481	0.2385	0.1089	0.0373	0.0102	0.0023

II) Piezas defectuosas. Una máquina fabrica tornillos y por término medio 15 de cada mil son defectuosos. Estamos interesados en saber, ¿cuántos tornillos deben ponerse en una caja para que la probabilidad de encontrar al menos 100 tornillos correctos sea mayor o igual que 0.8?.

En el supuesto de presentarlos en cajas de 100 unidades la probabilidad de no encontrar tornillos defectuosos en una caja puede evaluarse a partir de una Binomial $B(100, 0.015)$, de manera que

$$P(X=0) = (0.985)^{100} = 0.22061$$

o bien mediante una aproximación de Poisson, $P(1.5)$, $P(X=0) = e^{-1.5} = 0.22313\ldots$

aproximación que daremos por buena a efectos prácticos.

A la vista de las probabilidades obtenidas parece claro que las cajas deberán contener más de 100 tornillos, ~~[illegible]~~. Supongamos pues que la caja contiene $100 + x$ tornillos, con x entero y pequeño. Aplicamos la aproximación de Poisson con $\lambda = n.p = (100+x) \cdot 0.015 \simeq 1.5$. La probabilidad de que haya al menos 100 tornillos buenos es la misma de encontrar a lo sumo x tornillos defectuosos, por tanto

$$P(X \leq x) = \sum_{k=0}^{x} \frac{\lambda^k}{k!} \cdot e^{-\lambda} = e^{-1.5}\left[1 + \frac{1.5}{1} + \dots + \frac{(1.5)^x}{x!}\right] \geq 0.8$$

Existen tablas adecuadas para determinar el valor del sumatorio a partir del valor de x, encontramos en ellas que para $x=1$, $P(X \leq 1) \simeq 0.56$ y para $x=2$, $P(X \leq 2) = 0.809$. Habrá pues que embalar los tornillos en cajas de 102 unidades.

El verdadero valor de $P(X \leq 2)$ obtenido a partir de una Binomial $B(102, 0.015)$ es $0.8022\ldots$ muy próximo al encontrado con nuestra aproximación.

III) Un ejemplo empírico. El número de veces que aparece el par (7,7) entre 100 pares de números aleatorios debe de seguir una distribución binomial $B(100, 0.01)$. En la tabla se muestran las frecuencias absolutas, N_k, de k apariciones del par (7,7) en 100 grupos de 100 pares de dígitos aleatorios. Se presentan también en la tabla los valores teóricos de la distribución binomial y de la correspondiente aproximación de Poisson, $P(\lambda)$, con $\lambda = n.p = 1$. Las frecuencias relativas observadas, $N_k/100$, están razonablemente de acuerdo con las probabilidades teóricas de ambas distribuciones discretas.

	$P(X=k)$		
k	$B(100, 0.01)$	$P(1)$	N_k
0	0.366032	0.367879	41
1	0.369730	0.367879	34
2	0.184865	0.183940	16
3	0.060999	0.061313	8
4	0.014942	0.015328	0
5	0.002898	0.003066	1
6	0.000463	0.000511	0
7	0.000063	0.000073	0
8	0.000007	0.000009	0
9	0.000001	0.000001	0