Distribuciones discretas

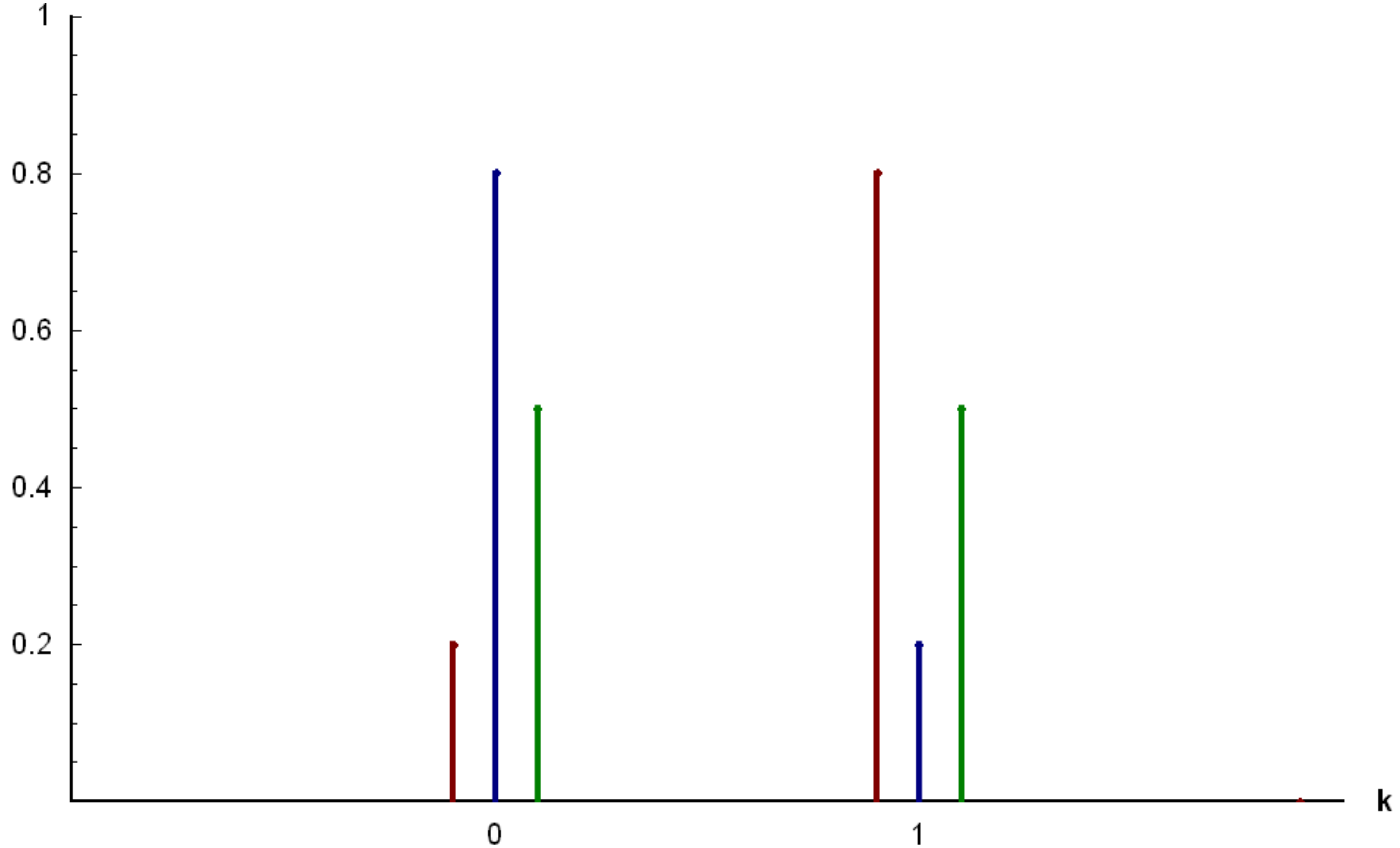
Sebastián Ruiz-Blais, I-2024

Contenidos

- Distribución de Bernoulli
- Distribución binomial
- Distribución multinomial
- Distribución hipergeométrica
- Distribución geométrica
- Distribución binomial negativa
- Distribución de Poisson

Distribución de Bernoulli

- Se realiza un ensayo, que se puede clasificar como éxito (1) o fracaso (0)
- La probabilidad de un **éxito** se denota como p y la de un **fracaso** es q



¿Qué observan en la relación entre *p* y *q*?

Problema

- Se seleccionan **3 artículos** al azar de un proceso de producción. Uno de cada 10 artículos sale defectuoso. ¿Cuál es la distribución de probabilidad de la variable aleatoria X, que determina el número de artículos defectuosos?

Proceso de Bernoulli

- Se realizan ensayos repetidos
- Cada ensayo se puede clasificar como éxito o fracaso
- La probabilidad de un éxito se denota como p y la de fracaso como q
- Los ensayos repetidos son independientes

Distribución binomial

Distribución binomial Un experimento de Bernoulli puede tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$. Entonces, la distribución de probabilidad de la variable aleatoria binomial X, el número de éxitos en n ensayos independientes, es

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Distribución de Bernoulli y Binomial

- La distribución de Bernoulli es un caso específico de la distribución Binomial, tomando $n=1$.
- La distribución Binomial se obtiene al repetir el experimento *n* veces

Problema

- La probabilidad de que un paciente se recupere de una rara enfermedad sanguínea es de 0.4. Si se sabe que 15 personas contrajeron la enfermedad, ¿cuál es la probabilidad de que a) sobrevivan al menos 10, b) sobrevivan de 3 a 8, y c) sobrevivan exactamente 5?

Solución

Solución: Sea X el número de personas que sobreviven.

a) $$P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} b(x; 15, 0.4) = 1 - 0.9662$$
$$= 0.0338$$

b) $$P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x; 15, 0.4) = \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4)$$
$$= 0.9050 - 0.0271 = 0.8779$$

c) $$P(X = 5) = b(5; 15, 0.4) = \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4)$$
$$= 0.4032 - 0.2173 = 0.1859$$

Ejercicio

La probabilidad de que cierta clase de componente sobreviva a una prueba de choque es de 3/4. Calcule la probabilidad de que sobrevivan exactamente 2 de los siguientes 4 componentes que se prueben.

Solución

***Solución*:** Si suponemos que las pruebas son independientes y $p = 3/4$ para cada una de las 4 pruebas, obtenemos

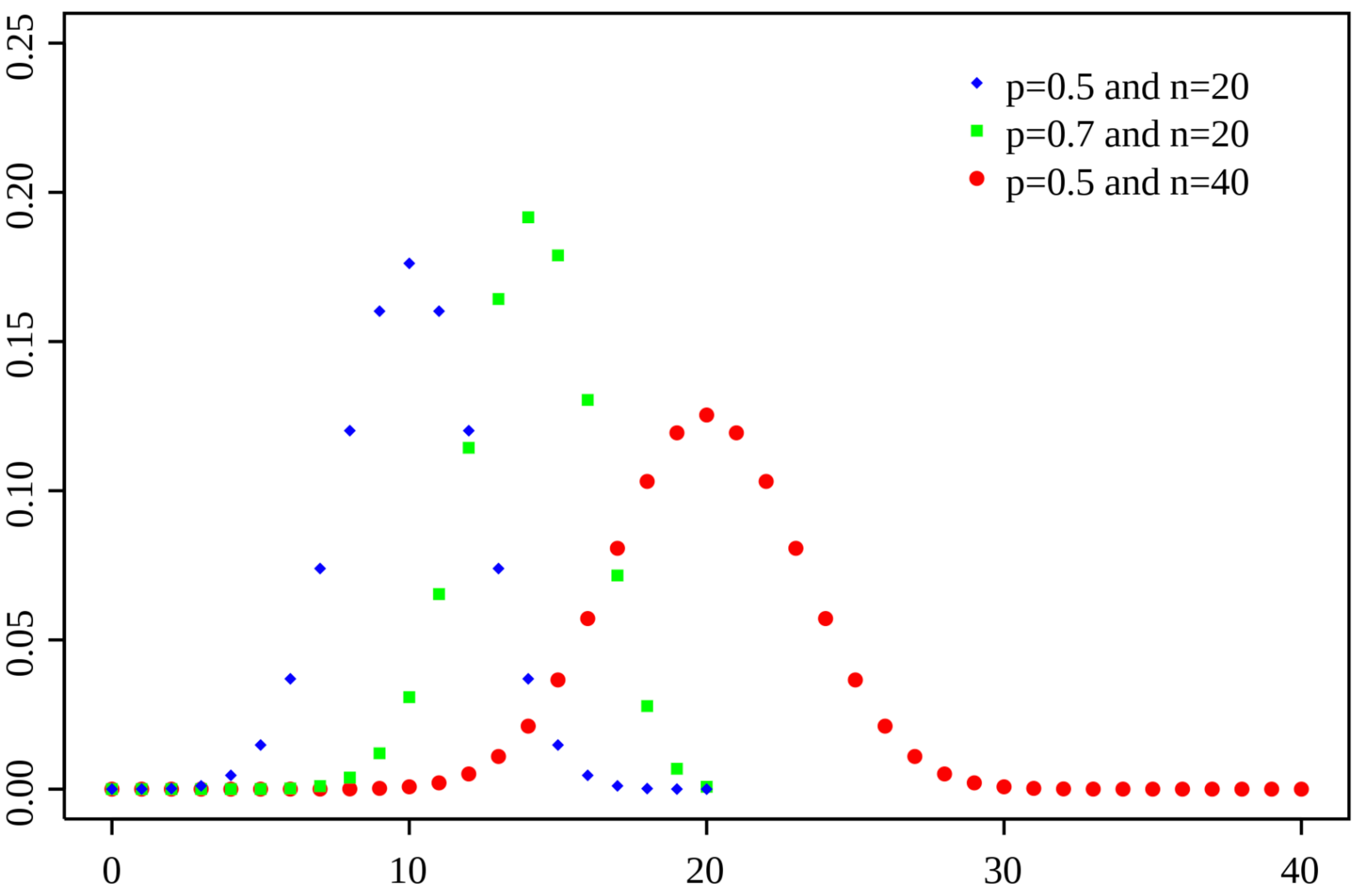
$$b\left(2; 4, \frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \left(\frac{4!}{2!\,2!}\right)\left(\frac{3^2}{4^4}\right) = \frac{27}{128}.$$

Media y varianza de la distribución binomial

Teorema 5.1: La media y la varianza de la distribución binomial $b(x; n, p)$ son

$$\mu = np \text{ y } \sigma^2 = npq.$$

Algunas distribuciones binomiales



¿Cuántos puntos tiene cada distribución? ¿Por qué?

¿Por qué difieren las distribuciones? Compruébelo

Problema

Ejemplo 5.7: La complejidad de las llegadas y las salidas de los aviones en un aeropuerto es tal que a menudo se utiliza la simulación por computadora para modelar las condiciones "ideales". Para un aeropuerto específico que tiene tres pistas se sabe que, en el escenario ideal, las probabilidades de que las pistas individuales sean utilizadas por un avión comercial que llega aleatoriamente son las siguientes:

Pista 1: $p_1 = 2/9$
Pista 2: $p_2 = 1/6$
Pista 3: $p_3 = 11/18$

¿Cuál es la probabilidad de que 6 aviones que llegan al azar se distribuyan de la siguiente manera?

Pista 1: 2 aviones
Pista 2: 1 avión
Pista 3: 3 aviones

Solución

***Solución*:** Si usamos la distribución multinomial, tenemos

$$f\left(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6\right) = \binom{6}{2, 1, 3}\left(\frac{2}{9}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{11}{18}\right)^3$$

$$= \frac{6!}{2!1!3!} \cdot \frac{2^2}{9^2} \cdot \frac{1}{6} \cdot \frac{11^3}{18^3} = 0.1127.$$

Distribución multinomial

Distribución multinomial Si un ensayo dado puede producir los k resultados $E_1, E_2, \ldots, E_k$ con probabilidades $p_1, p_2, \ldots, p_k$, entonces la distribución de probabilidad de las variables aleatorias $X_1, X_2, \ldots, X_k$, que representa el número de ocurrencias para $E_1, E_2, \ldots, E_k$ en n ensayos independientes, es

$$f(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},$$

con

$$\sum_{i=1}^{k} x_i = n \text{ y } \sum_{i=1}^{k} p_i = 1.$$

Problema

- Se tiene una baraja de 52 naipes. Si se sacan 5 naipes al azar (sin reemplazo), **¿cuál es la probabilidad de obtener 3 cartas rojas y 2 cartas negras?**

Solución

$$\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = \frac{(26!/3!23!)(26!/2!24!)}{52!/5!47!} = 0.3251.$$

Distribución hipergeométrica

Distribución hipergeométrica

La distribución de probabilidad de la variable aleatoria hipergeométrica X, el número de éxitos en una muestra aleatoria de tamaño n que se selecciona de N artículos, en los que k se denomina **éxito** y $N - k$ **fracaso**, es

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad \text{máx}\{0, n-(N-k)\} \le x \le \text{mín}\{n, k\}.$$

Problema

- Lotes con 40 componentes cada uno que contengan 3 o más defectuosos se consideran inaceptables. El procedimiento para obtener muestras del lote consiste en seleccionar una muestra de 5 componentes al azar y rechazar el lote si se encuentra un componente defectuoso. **¿Cuál es la probabilidad de que en la muestra se encuentre exactamente un componente defectuoso, si en todo el lote hay 3 defectuosos?**

Solución

Si utilizamos la distribución hipergeométrica con $n = 5$, $N = 40$, $k = 3$ y $x = 1$, encontramos que la probabilidad de obtener un componente defectuoso es

$$h(1; 40, 5, 3) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3011.$$

De nueva cuenta este plan no es adecuado porque sólo 30% de las veces detecta un lote malo (con 3 componentes defectuosos). ■

Problema

- Lotes con 40 componentes cada uno que contengan 3 o más defectuosos se consideran inaceptables. El procedimiento para obtener muestras del lote consiste en seleccionar una muestra de 5 componentes al azar y rechazar el lote si se encuentra un componente defectuoso.
- Para este problema, **¿cuál es la probabilidad de que en la muestra haya por lo menos un componente defectuoso, dado que el lote tiene 3 o más defectuosos?**

Media y varianza de distribución hipergeométrica

Teorema 5.2: La media y la varianza de la distribución hipergeométrica $h(x; N, n, k)$ son

$$\mu = \frac{nk}{N} \text{ y } \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N}\left(1 - \frac{k}{N}\right).$$

Problema

- Se usa un grupo de 10 individuos para un estudio de caso biológico. El grupo contiene 3 personas con sangre tipo O, 4 con sangre tipo A y 3 con tipo B. **¿Cuál es la probabilidad de que una muestra aleatoria de 5 contenga 1 persona con sangre tipo O, 2 personas con tipo A y 2 personas con tipo B?**

Solución

Si se utiliza la extensión de la distribución hipergeométrica con $x_1 = 1$, $x_2 = 2$, $x_3 = 2$, $a_1 = 3$, $a_2 = 4$, $a_3 = 3$, $N = 10$ y $n = 5$, vemos que la probabilidad que se desea es

$$f(1, 2, 2; 3, 4, 3, 10, 5) = \frac{\binom{3}{1}\binom{4}{2}\binom{3}{2}}{\binom{10}{5}} = \frac{3}{14}.$$

Distribución hipergeométrica multivariada

Distribución hipergeométrica multivariada

Si N artículos se pueden dividir en las k celdas $A_1, A_2, \ldots, A_k$ con $a_1, a_2, \ldots, a_k$ elementos, respectivamente, entonces la distribución de probabilidad de las variables aleatorias $X_1, X_2, \ldots, X_k$, que representan el número de elementos que se seleccionan de $A_1, A_2, \ldots, A_k$ en una muestra aleatoria de tamaño n, es

$$f(x_1, x_2, \ldots, x_k; a_1, a_2, \ldots, a_k, N, n) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}\cdots\binom{a_k}{x_k}}{\binom{N}{n}},$$

con $\sum_{i=1}^{k} x_i = n$ y $\sum_{i=1}^{k} a_i = N$.

Problema

- En la serie de campeonato de la NBA (National Basketball Association), el equipo que gane 4 de 7 juegos será el ganador. Suponga que los equipos A y B se enfrentan en los juegos de campeonato y que el equipo A tiene una probabilidad de 0.55 de ganarle al equipo B.
- a) ¿Cuál es la probabilidad de que el equipo A gane la serie en 6 juegos?
- b) ¿Cuál es la probabilidad de que el equipo A gane la serie?
- c) Si ambos equipos se enfrentaran en la eliminatoria de una serie regional y el triunfador fuera el que ganara 3 de 5 juegos, ¿cuál es la probabilidad de que el equipo A gane la serie?

Solución

a) $b^*(6;4,0.55) = \binom{5}{3}0.55^4(1-0.55)^{6-4} = 0.1853.$

b) P(el equipo A gana la serie de campeonato) es

$$\begin{aligned} &b^*(4;4,0.55) + b^*(5;4,0.55) + b^*(6;4,0.55) + b^*(7;4,0.55) \\ &\quad = 0.0915 + 0.1647 + 0.1853 + 0.1668 = 0.6083. \end{aligned}$$

c) P(el equipo A gana la eliminatoria) es

$$\begin{aligned} &b^*(3;3,0.55) + b^*(4;3,0.55) + b^*(5;3,0.55) \\ &\quad = 0.1664 + 0.2246 + 0.2021 = 0.5931. \end{aligned}$$

Distribución binomial negativa

Distribución binomial negativa Si ensayos independientes repetidos pueden dar como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X, el número del ensayo en el que ocurre el k-ésimo éxito, es

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \ldots$$

Problema

- Se sabe que en cierto proceso de fabricación uno de cada 100 artículos, en promedio, resulta defectuoso. **¿Cuál es la probabilidad de que el quinto artículo que se inspecciona, en un grupo de 100, sea el primer defectuoso que se encuentra?**

Distribución geométrica

Distribución geométrica Si pruebas independientes repetidas pueden tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X, el número de la prueba en el que ocurre el primer éxito, es

$$g(x; p) = pq^{x-1}, \quad x = 1, 2, 3, \ldots$$

Media y varianza para la distribución geométrica

Teorema 5.3: La media y la varianza de una variable aleatoria que sigue la distribución geométrica son

$$\mu = \frac{1}{p} \text{ y } \sigma^2 = \frac{1-p}{p^2}.$$

Problema

- En cierta fábrica los accidentes ocurren con muy poca frecuencia. Se sabe que la probabilidad de un accidente en cualquier día dado es de 0.005, y que los accidentes son independientes entre sí.
- a) **¿Cuál es la probabilidad de que ocurra un accidente en un día en un periodo de 400 días?**
- b) **¿Cuál es la probabilidad de que ocurra un accidente en tres días o menos de tal periodo?**

Problema

- Si bien este problema se puede resolver usando la distribución binomial, para simplificar los cálculos para números muy grandes se puede utilizar la aproximación de Poisson.

Solución

Sea X una variable aleatoria binomial con $n = 400$ y $p = 0.005$. Por consiguiente, $np = 2$. Si utilizamos la aproximación de Poisson,

a) $P(X = 1) = e^{-2}2^1 = 0.271$ y

b) $P(X \leq 3) = \sum_{x=0}^{3} e^{-2}2^x/x! = 0.857$.

Distribución de Poisson

Distribución de Poisson La distribución de probabilidad de la variable aleatoria de Poisson X, la cual representa el número de resultados que ocurren en un intervalo de tiempo dado o región específicos y se denota con t, es

$$p(x;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots,$$

donde λ es el número promedio de resultados por unidad de tiempo, distancia, área o volumen y $e = 2.71828\ldots$

Distribución de Poisson

- En lugar de hacer los cálculos con la fórmula de Poisson, también se puede utilizar las tablas que siguen al final de esta presentación.

Problema

- El número promedio de camiones-tanque que llega cada día a cierta ciudad portuaria es 10. Las instalaciones en el puerto pueden alojar un máximo de 15 camiones-tanque por día.
- **¿Cuál es la probabilidad de que en un día determinado lleguen más de 15 camiones y que por tanto se tenga que rechazar algunos?**

Solución

Sea X el número de camiones-tanque que llegan cada día. Entonces, usando la tabla A.2, tenemos

$$P(X > 15) = 1 - P(X \le 15) = 1 - \sum_{x=0}^{15} p(x; 10) = 1 - 0.9513 = 0.0487.$$

Aproximación de la binomial con Poisson

Teorema 5.5: Sea X una variable aleatoria binomial con distribución de probabilidad $b(x; n, p)$. Cuando $n \to \infty$, $p \to 0$, y $np \xrightarrow{n\to\infty} \mu$ permanece constante,

$$b(x; n, p) \xrightarrow{n\to\infty} p(x; \mu).$$

Media y varianza de la distribución Poisson

Teorema 5.4: Tanto la media como la varianza de la distribución de Poisson $p(x; \lambda t)$ son λt.

Tablas de Poisson

Sumas de probabilidad

Tabla A.2 Sumas de probabilidad de Poisson $\sum_{x=0}^{r} p(x; \mu)$

	μ								
r	**0.1**	**0.2**	**0.30**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

	μ								
r	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

Tabla A.2 (*continuación*) Sumas de probabilidad de Poisson $\sum_{x=0}^{r} p(x; \mu)$

	μ								
r	**5.5**	**6.0**	**6.5**	**7.0**	**7.5**	**8.0**	**8.5**	**9.0**	**9.5**
0	0.0041	0.0025	0.0015	0.0009	0.0006	0.0003	0.0002	0.0001	0.0001
1	0.0266	0.0174	0.0113	0.0073	0.0047	0.0030	0.0019	0.0012	0.0008
2	0.0884	0.0620	0.0430	0.0296	0.0203	0.0138	0.0093	0.0062	0.0042
3	0.2017	0.1512	0.1118	0.0818	0.0591	0.0424	0.0301	0.0212	0.0149
4	0.3575	0.2851	0.2237	0.1730	0.1321	0.0996	0.0744	0.0550	0.0403
5	0.5289	0.4457	0.3690	0.3007	0.2414	0.1912	0.1496	0.1157	0.0885
6	0.6860	0.6063	0.5265	0.4497	0.3782	0.3134	0.2562	0.2068	0.1649
7	0.8095	0.7440	0.6728	0.5987	0.5246	0.4530	0.3856	0.3239	0.2687
8	0.8944	0.8472	0.7916	0.7291	0.6620	0.5925	0.5231	0.4557	0.3918
9	0.9462	0.9161	0.8774	0.8305	0.7764	0.7166	0.6530	0.5874	0.5218
10	0.9747	0.9574	0.9332	0.9015	0.8622	0.8159	0.7634	0.7060	0.6453
11	0.9890	0.9799	0.9661	0.9467	0.9208	0.8881	0.8487	0.8030	0.7520
12	0.9955	0.9912	0.9840	0.9730	0.9573	0.9362	0.9091	0.8758	0.8364
13	0.9983	0.9964	0.9929	0.9872	0.9784	0.9658	0.9486	0.9261	0.8981
14	0.9994	0.9986	0.9970	0.9943	0.9897	0.9827	0.9726	0.9585	0.9400
15	0.9998	0.9995	0.9988	0.9976	0.9954	0.9918	0.9862	0.9780	0.9665
16	0.9999	0.9998	0.9996	0.9990	0.9980	0.9963	0.9934	0.9889	0.9823
17	1.0000	0.9999	0.9998	0.9996	0.9992	0.9984	0.9970	0.9947	0.9911
18		1.0000	0.9999	0.9999	0.9997	0.9993	0.9987	0.9976	0.9957
19			1.0000	1.0000	0.9999	0.9997	0.9995	0.9989	0.9980
20						0.9999	0.9998	0.9996	0.9991
21						1.0000	0.9999	0.9998	0.9996
22							1.0000	0.9999	0.9999
23								1.0000	0.9999
24									1.0000

Tabla A.2 (*continuación*) Sumas de probabilidad de Poisson $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	**10.0**	**11.0**	**12.0**	**13.0**	**14.0**	**15.0**	**16.0**	**17.0**	**18.0**
0	0.0000	0.0000	0.0000						
1	0.0005	0.0002	0.0001	0.0000	0.0000				
2	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000		
3	0.0103	0.0049	0.0023	0.0011	0.0005	0.0002	0.0001	0.0000	0.0000
4	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001
5	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003
6	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010
7	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029
8	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071
9	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154
10	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304
11	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549
12	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917
13	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426
14	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081
15	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867
16	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3751
17	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686
18	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622
19	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509
20	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307
21	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615	0.7991
22	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047	0.8551
23	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367	0.8989
24	1.0000	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594	0.9317
25		0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748	0.9554
26		1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848	0.9718
27			0.9999	0.9998	0.9994	0.9983	0.9959	0.9912	0.9827
28			1.0000	0.9999	0.9997	0.9991	0.9978	0.9950	0.9897
29				1.0000	0.9999	0.9996	0.9989	0.9973	0.9941
30					0.9999	0.9998	0.9994	0.9986	0.9967
31					1.0000	0.9999	0.9997	0.9993	0.9982
32						1.0000	0.9999	0.9996	0.9990
33							0.9999	0.9998	0.9995
34							1.0000	0.9999	0.9998
35								1.0000	0.9999
36									0.9999
37									1.0000