Probabilidad y Estadística (93.24)
Variables aleatorias continuas

Índice

1. Repaso de algunos conceptos

Variable aleatoria continua.

Una variable aleatoria real X es continua si

$$\mathrm{P}\left(X=x\right)=0 \quad \forall x \in \mathbb{R} \ \text{ y } \ \mathrm{P}\left(X \in \mathbb{R}\right)=1. \tag{1}$$

El comportamiento de una variable aleatoria continua (v.a.c.) está determinado por su *función de distribución de probabilidad acumulada* definida como

$$\mathrm{F}_X(x)=\mathrm{P}\left(X \leq x\right) \qquad \forall x \in \mathbb{R}. \tag{2}$$

En particular, la probabilidad de que $X \in (a, b]$ viene dada por

$$\mathrm{P}\left(a < X \leq b\right)=\mathrm{F}_X(b)-\mathrm{F}_X(a). \tag{3}$$

Es fácil mostrar que F_X cumple siempre las siguientes condiciones:

- Es continua (porque la probabilidad de un punto es nula).
- Es monótona no decreciente (porque la probabilidad es siempre no-negativa).
- Dado que X toma valores reales,

$$\mathrm{F}_X(-\infty)=\lim_{x \to -\infty} \mathrm{F}_X(x)=0, \qquad \mathrm{F}_X(+\infty)=\lim_{x \to +\infty} \mathrm{F}_X(x)=1. \tag{4}$$

También se suele asociar a cada v.a.c. una *densidad de probabilidad* definida como

$$\mathrm{f}_X(x)=\frac{d\mathrm{F}_X}{dx} \qquad \text{para todo } x \text{ donde la derivada existe.} \tag{5}$$

Por la monotonía de la función acumulada, f_X es siempre no-negativa. F_X puede ser calculada a partir de la densidad como

$$\mathrm{F}_X(x)=\int_{-\infty}^{x} \mathrm{f}_X(y)dy. \tag{6}$$

Obsérvese que $\int_{-\infty}^{+\infty} \mathrm{f}_X(y)dy=1$.

Valores esperados de una v.a.c.

Se define el *valor esperado de una función* $g(\cdot)$ *de una variable aleatoria continua* como

$$\mathrm{E}\left[g(X)\right] = \int_{-\infty}^{+\infty} g(x)\mathrm{f}_X(x)dx, \tag{7}$$

siempre que la integral en el lado de la derecha exista. En particular, los *momentos* de una v.a.c. vienen definidos por

$$\mathrm{E}\left[X^k\right] = \int_{-\infty}^{+\infty} x^k\mathrm{f}_X(x)dx \qquad \forall k \in \mathbb{N}. \tag{8}$$

Cuando existe $\mathrm{E}\left[X\right]$, también se definen los *momentos centrados* como

$$\mathrm{E}\left[(X-\mathrm{E}\left[X\right])^k\right] = \int_{-\infty}^{+\infty} (x-\mathrm{E}\left[X\right])^k\mathrm{f}_X(x)dx \qquad \forall k \in \mathbb{N}. \tag{9}$$

Dos parámetros importantes de toda v.a.c. son su *valor esperado* o *media* y su *varianza* dados por

$$\mu_X = \mathrm{E}\left[X\right] = \int_{-\infty}^{+\infty} x\mathrm{f}_X(x)dx, \tag{10}$$

$$\sigma_X^2 = \mathrm{Var}\left[X\right] = \mathrm{E}\left[(X-\mu_X)^2\right] = \mathrm{E}\left[X^2\right] - \left(\mathrm{E}\left[X\right]\right)^2, \tag{11}$$

respectivamente.

Integrales impropias

La integral impropia converge si el límite que lo define existe.

$$\int_{a}^{+\infty} w(x)dx = \lim_{t\to+\infty}\int_{a}^{t} w(x)dx, \tag{12}$$

$$\int_{-\infty}^{b} w(x)dx = \lim_{t\to-\infty}\int_{t}^{a} w(x)dx, \tag{13}$$

$$\int_{a}^{b} w(x)dx = \lim_{t\to c^-}\int_{a}^{t} w(x)dx + \lim_{t\to c^+}\int_{t}^{b} w(x)dx, \tag{14}$$

donde el último caso corresponde a cuando $w(x)$ no es continua o no está definida en c.

Distribución Uniforme

Se dice que una variable aleatoria continua X tiene distribución uniforme entre a y b $(a < b)$ si su densidad de probabilidad viene dada por

$$\mathrm{f}_X(x) = \begin{cases} 0 & \text{si } x \notin (a, b), \\ \frac{1}{b-a} & \text{si } x \in (a, b). \end{cases} \tag{15}$$

La función de probabilidad acumulada es

$$\mathrm{F}_X(x) = \begin{cases} 0 & \text{si } x \leq a, \\ \frac{x-a}{b-a} & \text{si } x \in (a, b), \\ 1 & \text{si } x \geq b. \end{cases} \tag{16}$$

Es fácil calcular

$$\mu_X = \mathrm{E}\left[X\right] = \frac{a+b}{2}, \tag{17}$$

$$\sigma_X^2 = \mathrm{Var}\left[X\right] = \frac{(b-a)^2}{12}. \tag{18}$$

Distribución exponencial

Se dice que una variable aleatoria continua T tiene distribución exponencial con parámetro λ si su densidad de probabilidad viene dada por

$$f_T(t) = \begin{cases} 0 & \text{si } t < 0, \\ \lambda \exp\left\{-\lambda t\right\} & \text{si } t > 0. \end{cases} \tag{19}$$

La función de probabilidad acumulada es

$$F_T(t) = \begin{cases} 0 & \text{si } t < 0, \\ 1 - \exp\left\{-\lambda t\right\} & \text{si } t > 0. \end{cases} \tag{20}$$

Es fácil calcular

$$\mu_T = \mathrm{E}\left[T\right] = \frac{1}{\lambda}, \tag{21}$$

$$\sigma_T^2 = \mathrm{Var}\left[T\right] = \frac{1}{\lambda^2}. \tag{22}$$

Comúnmente, las variables aleatorias exponenciales son utilizadas para modelar intervalos de tiempo con ciertas características particulares. Una de estas características es la *ausencia de memoria* de una v.a. exponencial, esto es, el hecho que

$$\mathrm{P}\big(T \geq t + \Delta t \big| \, T \geq t\big) = \mathrm{P}\left(T \geq \Delta t\right). \tag{23}$$

Distribución normal o Gaussiana

Se dice que una variable aleatoria continua X tiene distribución normal o Gaussiana con media μ_X y desvío estándar σ_X, $X \sim \mathcal{N}(\mu_X, \sigma_X)$ si su densidad de probabilidad viene dada por

$$\mathrm{f}_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right\}. \tag{24}$$

La densidad de probabilidad es simétrica respecto de $x = \mu_X$. Por ello, la mediana coincide con el valor esperado o media.
No existe expresión analítica de la función de probabilidad acumulada. Sin embargo, se utilizan datos calculados numéricamente y tabulados para una variable aleatoria $Z \sim \mathcal{N}(0,1)$:

$$\Phi(z) = \mathrm{P}(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy. \tag{25}$$

Luego, es fácil verificar que la función de probabilidad acumulada de X (cualquier media y varianza dados) es

$$\mathrm{F}_X(x) = \mathrm{P}(X \le x) = \mathrm{P}\left(\frac{X-\mu_X}{\sigma_X} \le \frac{x-\mu_X}{\sigma_X}\right) = \Phi\left(\frac{x-\mu_X}{\sigma_X}\right). \tag{26}$$

En el manejo de la función $\Phi(z)$ es útil recordar los siguientes hechos basados en la simetría de la densidad de probabilidad:

- $\Phi(-z) = 1 - \Phi(z)\ \forall z \in \mathbb{R}$.
- $\Phi^{-1}(\alpha) = -\Phi^{-1}(1-\alpha)\ \forall \alpha \in (0,1)$.

Finalmente, es habitual que usemos la notación

$$z_\alpha = \Phi^{-1}(\alpha). \tag{27}$$

2. Función de distribución de la variable aleatoria con distribución normal estándar

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

3. Fractiles de la distribución normal estándar

α	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.50	0.0000	0.0025	0.0050	0.0075	0.0100	0.0125	0.0150	0.0175	0.0201	0.0226
0.51	0.0251	0.0276	0.0301	0.0326	0.0351	0.0376	0.0401	0.0426	0.0451	0.0476
0.52	0.0502	0.0527	0.0552	0.0577	0.0602	0.0627	0.0652	0.0677	0.0702	0.0728
0.53	0.0753	0.0778	0.0803	0.0828	0.0853	0.0878	0.0904	0.0929	0.0954	0.0979
0.54	0.1004	0.1030	0.1055	0.1080	0.1105	0.1130	0.1156	0.1181	0.1206	0.1231
0.55	0.1257	0.1282	0.1307	0.1332	0.1358	0.1383	0.1408	0.1434	0.1459	0.1484
0.56	0.1510	0.1535	0.1560	0.1586	0.1611	0.1637	0.1662	0.1687	0.1713	0.1738
0.57	0.1764	0.1789	0.1815	0.1840	0.1866	0.1891	0.1917	0.1942	0.1968	0.1993
0.58	0.2019	0.2045	0.2070	0.2096	0.2121	0.2147	0.2173	0.2198	0.2224	0.2250
0.59	0.2275	0.2301	0.2327	0.2353	0.2378	0.2404	0.2430	0.2456	0.2482	0.2508
0.60	0.2533	0.2559	0.2585	0.2611	0.2637	0.2663	0.2689	0.2715	0.2741	0.2767
0.61	0.2793	0.2819	0.2845	0.2871	0.2898	0.2924	0.2950	0.2976	0.3002	0.3029
0.62	0.3055	0.3081	0.3107	0.3134	0.3160	0.3186	0.3213	0.3239	0.3266	0.3292
0.63	0.3319	0.3345	0.3372	0.3398	0.3425	0.3451	0.3478	0.3505	0.3531	0.3558
0.64	0.3585	0.3611	0.3638	0.3665	0.3692	0.3719	0.3745	0.3772	0.3799	0.3826
0.65	0.3853	0.3880	0.3907	0.3934	0.3961	0.3989	0.4016	0.4043	0.4070	0.4097
0.66	0.4125	0.4152	0.4179	0.4207	0.4234	0.4261	0.4289	0.4316	0.4344	0.4372
0.67	0.4399	0.4427	0.4454	0.4482	0.4510	0.4538	0.4565	0.4593	0.4621	0.4649
0.68	0.4677	0.4705	0.4733	0.4761	0.4789	0.4817	0.4845	0.4874	0.4902	0.4930
0.69	0.4958	0.4987	0.5015	0.5044	0.5072	0.5101	0.5129	0.5158	0.5187	0.5215
0.70	0.5244	0.5273	0.5302	0.5330	0.5359	0.5388	0.5417	0.5446	0.5476	0.5505
0.71	0.5534	0.5563	0.5592	0.5622	0.5651	0.5681	0.5710	0.5740	0.5769	0.5799
0.72	0.5828	0.5858	0.5888	0.5918	0.5948	0.5978	0.6008	0.6038	0.6068	0.6098
0.73	0.6128	0.6158	0.6189	0.6219	0.6250	0.6280	0.6311	0.6341	0.6372	0.6403
0.74	0.6433	0.6464	0.6495	0.6526	0.6557	0.6588	0.6620	0.6651	0.6682	0.6713
0.75	0.6745	0.6776	0.6808	0.6840	0.6871	0.6903	0.6935	0.6967	0.6999	0.7031
0.76	0.7063	0.7095	0.7128	0.7160	0.7192	0.7225	0.7257	0.7290	0.7323	0.7356
0.77	0.7388	0.7421	0.7454	0.7488	0.7521	0.7554	0.7588	0.7621	0.7655	0.7688
0.78	0.7722	0.7756	0.7790	0.7824	0.7858	0.7892	0.7926	0.7961	0.7995	0.8030
0.79	0.8064	0.8099	0.8134	0.8169	0.8204	0.8239	0.8274	0.8310	0.8345	0.8381
0.80	0.8416	0.8452	0.8488	0.8524	0.8560	0.8596	0.8632	0.8669	0.8706	0.8742
0.81	0.8779	0.8816	0.8853	0.8890	0.8927	0.8965	0.9002	0.9040	0.9078	0.9116
0.82	0.9154	0.9192	0.9230	0.9269	0.9307	0.9346	0.9385	0.9424	0.9463	0.9502
0.83	0.9542	0.9581	0.9621	0.9661	0.9701	0.9741	0.9782	0.9822	0.9863	0.9904
0.84	0.9945	0.9986	1.0027	1.0069	1.0110	1.0152	1.0194	1.0237	1.0279	1.0322
0.85	1.0364	1.0407	1.0451	1.0494	1.0537	1.0581	1.0625	1.0669	1.0714	1.0758
0.86	1.0803	1.0848	1.0893	1.0939	1.0985	1.1031	1.1077	1.1123	1.1170	1.1217
0.87	1.1264	1.1311	1.1359	1.1407	1.1455	1.1503	1.1552	1.1601	1.1650	1.1700
0.88	1.1750	1.1800	1.1850	1.1901	1.1952	1.2004	1.2055	1.2107	1.2160	1.2212
0.89	1.2265	1.2319	1.2372	1.2426	1.2481	1.2536	1.2591	1.2646	1.2702	1.2759
0.90	1.2816	1.2873	1.2930	1.2988	1.3047	1.3106	1.3165	1.3225	1.3285	1.3346
0.91	1.3408	1.3469	1.3532	1.3595	1.3658	1.3722	1.3787	1.3852	1.3917	1.3984
0.92	1.4051	1.4118	1.4187	1.4255	1.4325	1.4395	1.4466	1.4538	1.4611	1.4684
0.93	1.4758	1.4833	1.4909	1.4985	1.5063	1.5141	1.5220	1.5301	1.5382	1.5464
0.94	1.5548	1.5632	1.5718	1.5805	1.5893	1.5982	1.6072	1.6164	1.6258	1.6352
0.95	1.6449	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392
0.96	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663
0.97	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335
0.98	2.0537	2.0748	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904
0.99	2.3263	2.3656	2.4089	2.4573	2.5121	2.5758	2.6521	2.7478	2.8782	3.0902

4. Guía de ejercicios

1. Una variable aleatoria X tiene densidad de probabilidad

$$\mathrm{f}_X(x) = \begin{cases} \frac{c}{\sqrt{A^2-x^2}} & x \in (-A, A) \\ 0 & x \notin (-A, A) \end{cases}$$

   a) Determine el valor de la constante c para que sean funciones densidad de probabilidad de variables aleatorias continuas (variable aleatoria continua) X;
   b) Encuentre la función de distribución $\mathrm{F}_X(x) = \mathrm{P}(X \leq x)$.
   c) Calcule el valor esperado $\mathrm{E}[X]$ y la varianza $\mathrm{Var}[X]$;

2. Sea X una čon distribución uniforme en $(0, 8)$. Calcular:

   a) $\mathrm{P}(2 < X < 5)$.
   b) $\mathrm{P}\big(0 < X < 3 | 0 < X < 5\big)$.

3. La función densidad de probabilidad $\mathrm{f}_X(x)$ de la variable aleatoria continua X es:

$$\mathrm{f}_X(x) = \begin{cases} 2(1-x) & x \in (0, 1) \\ 0 & x \notin (0, 1) \end{cases}$$

   a) Calcular $\mathrm{E}[X]$ y $\mathrm{Var}[X]$.
   b) Calcular la mediana m de X definida por $\mathrm{P}(X < m) = 0.5$.

   Resolución aquí.

4. La variable aleatoria continua X tiene por función densidad de probabilidad a $\mathrm{f}_X(x) = x/2 \;\; 0 < x < 2$ y 0 fuera de ese intervalo. Si se hacen

   a) dos determinaciones independientes de X, ¿cuál es la probabilidad de que ambas sean mayores que 1?
   b) tres determinaciones independientes, ¿cuál es la probabilidad de que exactamente dos sean mayores que 1?

5. El porcentaje de alcohol $(100\ X)$ en cierto compuesto se puede considerar una variable aleatoria continua, en donde X verifica $0 < X < 1$ y tiene la función de densidad $\mathrm{f}_X(x) = 20x^3(1-x) \;\; 0 < x < 1$ y 0 fuera de ese intervalo.

   a) Determinar la función de distribución F_X y representarla gráficamente.
   b) Calcular $P[(X < 1/2)/(1/3 < X < 2/3)]$.

6. La duración en horas X de cierto dispositivo electrónico es una variable aleatoria continua con función densidad de probabilidad

$$\mathrm{f}_X(x) = \begin{cases} 100/x^2 & x > 100 \\ 0 & x \leq 100 \end{cases}$$

   a) ¿Cuál es la probabilidad de que un dispositivo dure menos de 200 hs si se sabe que el dispositivo aún funciona después de 150 hs de servicio?
   b) ¿Cuál es la probabilidad de que si se instalan 3 de tales dispositivos en un sistema 2 fallen antes de las 150 horas de servicio y el restante después de las 150 hs?

c) ¿Cuál es el número máximo de dispositivos que se pueden poner en un sistema para que la probabilidad de que después de 150 horas de servicio todos ellos funcionen sea 0.4?

Resolución aquí. Video con el ejercicio explicado.

7. El colectivo de cierta línea va a un horario estricto con intervalos de 5 minutos. Calcular la probabilidad de que un pasajero que se acerque a la parada en un instante al azar tenga que esperar el colectivo menos de tres minutos. Suponga que el instante de llegada del pasajero a la parada es una variable aleatoria continua con distribución uniforme en (0, 5).

8. La duración en horas de un sistema puede considerarse una variable aleatoria continua con distribución exponencial de parámetro característico 0.001 hora^{-1}.

   a) Una empresa que produce este sistema desea garantizarlo durante cierto tiempo. ¿Cuántas horas lo debe amparar la garantía de buen funcionamiento para asegurar, con probabilidad 0.95, que el sistema funcionará como mínimo el número de horas garantizadas.

   b) Un dispositivo utiliza cinco de estos sistemas y funciona si todos ellos lo hacen. Los sistemas funcionan (y fallan) en forma independiente. ¿Cuál es la probabilidad de que el dispositivo opere, por lo menos, 100 horas? ¿Y 1000 horas?

   c) Otro dispositivo también utiliza cinco de estos sistemas y funciona mientras lo hagan por lo menos tres de ellos (no suponga ninguna conexión en especial). Los sistemas funcionan (y también fallan) en forma independiente. ¿Cuál es la probabilidad de que el dispositivo funcione aún al cabo de 1000 horas?

9. Un sistema está integrado por dos componentes A y B que fallan al azar en promedio 1 vez cada 500 y 800 horas respectivamente. El sistema falla cuando cualquiera de dichos componentes falla. Se supone que estos componentes fallan en forma independiente y que el tiempo entre fallas es una variable aleatoria con distribución exponencial. ¿Cuál es la probabilidad de que el sistema falle después de transcurridas las primeras 1000 horas?

10. El tiempo de duración hasta la falla de un dispositivo mecánico se supone que tiene distribución exponencial con una media de 400 hs.

    a) Este dispositivo ha funcionado sin fallas durante 400 horas. ¿Cuál es la probabilidad de que falle en las próximas 100 horas?.

    b) Si se ponen a funcionar 10 de estos dispositivos , ¿cuál es la probabilidad de que falle al menos uno de ellos antes de 100 horas?. Suponga que los dispositivos fallan de manera independiente.

    Resolución aquí.

11. Un sistema consta de n componentes idénticos conectados en serie. Cuando falla por lo menos un componente falla todo el sistema. Suponga que la duración de cada componente es una variable aleatoria con distribución exponencial de parámetro $\lambda = 0.01$ 1/hora, y que los componentes fallan independientemente uno de otro. Defina los sucesos independientes A_i = i- ésimo componente dura por lo menos t horas, $i = 1, \ldots n$. Sea T el tiempo en el que falla el sistema, esto es, la duración mínima de funcionamiento entre los n componentes.
    Considerando la independencia de los sucesos A_i, obtenga $\mathrm{F}_T(t) = \mathrm{P}(T < t)$ y la función densidad de probabilidad de T. ¿Qué tipo de distribución de probabilidades tiene T ? ¿Cuál es el valor esperado de T?

12. La duración de ciertos dispositivos es una variable aleatoria con distribución exponencial de parámetro λ (en 1/hora). La probabilidad de que en una muestra de 5 dispositivos haya por lo menos uno que dure más de 1200 hs. es 0.75.

    a) Hallar la constante λ.

    b) Suponga que estos dispositivos se conectan en serie de manera tal que al fallar por lo menos uno de ellos falla todo el sistema. Calcular la probabilidad de que el sistema funcione más de 1500 horas.

    Resolución aquí. Video con el ejercicio explicado.

13. Una función que se relaciona a veces con variable aleatoria continua no negativas es la función de tasa de fallas (o la tasa de riesgo). Esta función se define por $R(t) = \mathrm{f}_T(t)/(1-\mathrm{F}_T(t))$ para una función de densidad $\mathrm{f}_T(t)$ con su correspondiente función de distribución $\mathrm{F}_T(t)$. Si la variable aleatoria T es la duración de un componente entonces, para Δt pequeño, $\mathrm{P}\big(t \leq T \leq t+\Delta t | T > t\big) \approx R(t)\Delta t$. Demostrar que una variable aleatoria continua que toma valores reales positivos tiene distribución exponencial si y solo si $R(t)$ es constante.

    Comentario: Observe que la condición es necesaria y suficiente, si T tiene distribución exponencial de parámetro λ usando la definición de $R(t)$ resulta que $R(t) = \lambda$. Con la condición $R(t) = \lambda$ se obtiene un problema de valor inicial para $\mathrm{F}_T(t)$, la función de distribución de T..

14. Se sabe que la duración X de una pieza responde a la distribución de *Weibull* cuya función de distribución es: $\mathrm{F}_X(x) = 1-\exp(-(\lambda\, x)^b)$ para $x > 0$ con b y λ constantes reales positivas. Si x se mide en miles de horas, $b = 2$ y $\lambda = 0.01$:

    a) obtener la función de densidad de probabilidad de la duración de la pieza.

    b) calcular la duración garantizada G_{90} con un 90 % de confiabilidad. *Nota:* Se cumple $\mathrm{P}\left(X > G_{90}\right) = 0.90$.

    c) Se realizó una prueba con un lote muy grande de piezas y se separaron 1000 con duración superior a 10000 horas; ¿cuántas unidades espera Ud. haya en este lote de 1000 cuyas duraciones totales sean superiores a 20000 horas?

    Resolución aquí. Video con el ejercicio explicado.

15. **Sobre la distribución normal**
    Suponga que la variable aleatoria continua X tiene distribución normal de media μ y desvío estándar σ. Sean f_X y F_X las funciones de densidad de probabilidad y de distribución respectivamente.

    a) Demostrar que $\mathrm{F}_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, donde Φ es la función de distribución de la variable continua Z que tiene distribución normal estándar (media cero y desvío estándar unitario).

    b) Demostrar que $\Phi(z) = 1 - \Phi(-z)$.

    c) Sea z_α el fractil α de la variable continua Z que tiene distribución normal estándar, entonces $\Phi(z_\alpha) = \alpha$. Demostrar que $z_{1-\omega} = -z_\omega$.

    d) Para una variable aleatoria X de media μ y desvío estándar σ el coeficiente de *kurtosis* κ viene dado por $\kappa = \frac{\mathrm{E}\left[(X-\mu)^4\right]}{\sigma^4}$. Demostrar que para la variable continua X que tiene distribución normal se verifica $\kappa = 3$. Es dato

$$\int_{-\infty}^{+\infty} z^4\, \exp(-a\, z^2)\, dz = \frac{3\sqrt{\pi}}{4}\, a^{-\frac{5}{2}}, \qquad a > 0.$$

e) Calcule los cuartiles primero Q_1, tercero Q_3 y el rango intercuartílico $I_Q = |Q_3 - Q_1|$ para la variable aleatoria X con distribución normal de media μ y desvío estándar σ. Según la definición de fractil se cumple $\mathrm{F}(Q_1) = 0.25$ mientras que $\mathrm{F}(Q_3) = 0.75$.

f) Según la convención introducida por J. W. Tukey para la construcción del *diagrama de caja* o *boxplot*, los valores medidos de la variable de interés en una muestra que sean menores que $q_1 - 1.5\, i_q$, o mayores que $q_3 + 1.5\, i_q$, son declarados *valores extremos* o *outliers* (q_1, q_3 e i_q son, respectivamente, los cuartiles primero, tercero y el rango intercuartílico para los datos de la muestra considerada). Calcule la probabilidad que una variable aleatoria X con distribución normal tome *valores extremos*.

g) Considere la variable continua Z con distribución normal estándar (media cero y desvío estándar unitario). Calcule $\mathrm{E}[|Z|]$.

16. El consumo mensual de combustible en una planta industrial se puede considerar una variable aleatoria continua con distribución normal con media 18000 litros y desvío estándar 1500 litros.

    a) ¿Qué porcentaje de los meses se consume menos de 21000 litros?

    b) ¿Qué porcentaje de los meses se consume más de 17000 litros?

    c) ¿Qué porcentaje de los meses el consumo es mayor que 17000 y menor que 21000 litros?.

    d) ¿Qué capacidad debe tener un tanque para satisfacer el consumo mensual con probabilidad 0.95?

    e) ¿Cuál es el consumo superado en el 90 % de los meses ?

    f) De los meses que se consume más de 17000 litros, ¿qué porcentaje se consume menos de 21000 litros ?

17. Por larga experiencia se ha determinado que la capacidad de cierto tipo de condensadores, fabricados con el mismo procedimiento, es una variable aleatoria continua con distribución normal de media 10 y desvío estándar 0.1 (ambos valores en microfaradios).

    a) Para que estos componentes puedan utilizarse en un sistema deben tener una capacidad no inferior a 9.8. ¿Cuál es el porcentaje de condensadores que no cumplen la especificación?

    b) ¿Para que valor de la capacidad C_0 el 99 % de los condensadores tienen una capacidad igual o menor que C_0?

18. Un comerciante tiene piezas de mármol en dos depósitos, 70 % y 30 % respectivamente, en cada uno de ellos. El peso de estas piezas en kg se distribuye así: $N_1(1200, 150)$ y $N_2(1100, 100)$, (una forma de anotar media y dispersión de una distribución normal). Por error se juntaron las piezas en un solo depósito y se debe enviar 50 de ellas con un peso no inferior a 1100 kg sin tiempo para revisarlas. Calcular la probabilidad de que más de 10 piezas sean devueltas.

19. Una máquina produce tornillos con un diámetro que se distribuye en forma normal con media 8 mm. Se sabe que el 10 % de la producción es frágil por tener un diámetro menor que 5 mm. ¿Qué porcentaje de la producción tiene un diámetro mayor que 10 mm?

20. Un fábrica entrega remaches cuyo diámetro tiene una distribución normal. Se sabe que el 2.81 % de los remaches entregados tiene un diámetro mayor que 2.1 mm y que el 13.79 % tiene un diámetro menor que 1.5 mm. ¿Cuál es el valor medio y la varianza del diámetro de los remaches entregados?

21. El diámetro de contacto de la rosca de una unión se distribuye normalmente con media 10.02 mm y desvío estándar 0.075 mm. Las especificaciones dadas para esa rosca son: 10.00 ± 0.025 mm (la especificación es un intervalo de valores dentro del cuál deberá estar el diámetro para que sea considerado aceptable).

    a) ¿Cuál es la probabilidad de que la rosca esté fuera de la especificación dada?

    b) Suponiendo que el diámetro de contacto esté centrado en el valor nominal de la especificación (10 mm), ¿cuál es la máxima desviación estándar del diámetro aceptable que permitirá no más de una defectuosa entre mil producidas?

    Resolución aquí. **Video con el ejercicio explicado.**

22. Dos máquinas rectificadoras A y B trabajan de tal manera que cuando se ajusta el diámetro requerido para la pieza rectifican según una distribución normal de media 5 mm y desvíos de 0.01 y 0.02 mm respectivamente. La máquina A hace el 80 % del trabajo y la B el resto. Se mezclan las piezas rectificadas y se supone que una pieza es buena si su diámetro D está comprendido entre 5 mm y 5.02 mm.

    a) ¿Cuál es la probabilidad de que una pieza extraída al azar resulte fuera de especificación?

    b) Si una pieza extraída al azar resulta fuera de especificación, ¿cuál es la probabilidad de que haya sido rectificada por la máquina B?

    c) Si se toma una muestra de 10 artículos de la mezcla de los fabricados por ambas máquinas ¿cuál es la probabilidad de encontrar i.) una pieza buena, ii.) por lo menos una pieza buena?

23. La vida promedio de cierto tipo de motor pequeño es de 10 años con una desviación típica de 2 años. El fabricante garantiza reponer todos los motores que fallen dentro de un cierto período. Suponiendo que la vida de un motor sea una variable aleatoria continua distribuida normalmente, hallar el período de garantía que debe ofrecer el fabricante si sólo desea tener que reponer el 3 % de los motores que fallen.

24. El gerente de producción de una fábrica estima que la vida útil de una máquina es una variable aleatoria continua distribuida normalmente con una media de 3000 horas. Si además considera que hay una probabilidad de 0.5 de que la máquina dure menos de 2632 o más de 3368 horas, ¿cuál es la desviación típica de la duración de dicha máquina?

25. Si el tiempo T para fallar de un sistema está distribuido normalmente con media 90 horas y un desvío estándar 5 horas, ¿cuántas horas de operación (t_c) deben considerarse para que la confiabilidad (probabilidad de que el tiempo de buen funcionamiento exceda t_c) sea 0.90? ¿ y 0.95?

26. Se puede ajustar una máquina expendedora de gaseosas de tal manera que llene los vasos con un promedio de m unidades de volumen (U_V) por vaso. Si se supone que el volumen despachado por la máquina es una variable aleatoria continua con distribución normal de media m y desvío estándar 0.3 U_V, determine el valor de m de manera que los vasos de 8 U_V solamente se derramen el 1 % de las veces en que son llenados.

27. Se supone que los resultados de un examen tienen una distribución normal con una media de 78 y una varianza de 36.

    a) Determinar la probabilidad de que una persona que realiza el examen obtenga una nota mayor que 72.

b) Suponga que a los estudiantes que se encuentran en el 10 % de la parte superior de la distribución se les asigna una calificación A. ¿Cuál es la nota mínima que se debe obtener para tener calificación A?

c) ¿Cuál debe ser la mínima nota aprobatoria si el evaluador pretende que solamente el 28.1 % de los estudiantes apruebe?

d) Calcular, aproximadamente, la proporción de los estudiantes que tienen nota que exceden por lo menos en 5 puntos a la calificación reprobatoria del 25 % (calificaciones inferiores).

e) Si se sabe que la calificación de un estudiante es mayor que 72, ¿cuál es la probabilidad de que su calificación sea mayor que 84?

Resolución aquí.

28. Un empleado va al trabajo en subte el 70 % de las veces y en colectivo el 30 % restante. El tiempo que tarda cuando va en subte es una variable aleatoria normal con media 30 min. y dispersión 5 min, mientras que el tiempo que tarda cuando va en colectivo es una variable aleatoria normal con media 40 min. y dispersión 8 min.

a) Si un día tarda en llegar al trabajo más de 35 minutos, ¿cuál es la probabilidad de que haya viajado en subte?

b) ¿Cuál es la probabilidad de que en al menos dos de 10 días consecutivos tarde más de 35 minutos en llegar al trabajo? Indique que suposición permite el cálculo de esta probabilidad.

29. La longitud de ciertos dispositivos mecánicos tiene la siguiente tolerancia: 58 $\pm$ 0.21 mm. Se han recibido tres partidas de distintos proveedores de 200, 300 y 500 cajas de 6 unidades cada una, que si bien no difieren en su valor medio 58 mm presentan distintas dispersiones: 0.102 mm, 0.12 mm y 0.146 mm. Se supone que la longitud de esos dispositivos es una variable aleatoria con distribución normal. En el almacén se han mezclado las cajas. Si se elige una caja al azar y se comprueba que exactamente 4 de las piezas están dentro de los límites de tolerancia, ¿cuál es la probabilidad de que la caja haya sido entregada por el proveedor número 2?

Resolución aquí. Video con el ejercicio explicado.

30. Las precipitaciones (en mm) en un dado período de tiempo puede modelarse como una variable aleatoria X con distribución Gama, esto es, $X \sim \text{Gama}(\alpha, \beta)$ con $\alpha, \beta > 0$. La función de densidad de probabilidad es

$$\mathrm{f}_X(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0, \\ 0 & x < 0, \end{cases} \tag{28}$$

donde $\Gamma(\alpha)$ es la función Gama definida como

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt. \tag{29}$$

En particular, para $n \in \mathbb{N}$, $\Gamma(n+1) = n!$.

a) Calcule el valor esperado y la varianza de X.

b) Si $\alpha = 0.95$ y $\beta^{-1} = 10.9$ mm, estime el valor esperado de la precipitación. Datos tomados de la Fig. 1 en
Wilks, Daniel S. “Maximum Likelihood Estimation for the Gamma Distribution Using Data Containing Zeros”. Journal of Climate, Vol. 3, No. 12, pp. 1495–1501, 1990.

c) Para los mismos valores de los parámetros y con ayuda de un programa matemático, calcule la probabilidad de que la precipitación supero 50 mm.

31. La longitud (en bytes) de *posts* en un foro online puede considerarse una variable aleatoria L con distribución lognormal, esto es, $L \sim \text{LogNorm}(\mu, \sigma)$. La densidad de probabilidad es

$$f_L(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} \exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\} & x > 0, \\ 0 & x < 0. \end{cases} \tag{30}$$

a) Encuentre la función de distribución acumulada y el valor de la mediana.

b) Si $\mu = 5.417$ y $\sigma = 1.145$, encuentre la mediana. Estos valores fueron tomados de la Fig. 1 de

Sobkowicz, P., Thelwall, M., Buckley, K., Paltoglou, G. y Sobkowicz, A. "Lognormal distributions of user post lengths in Internet discussions-a consequence of the Weber-Fechner law?". EPJ Data Science, Vol. 2, pp.1-20, 2013.

c) Para los mismos valores de los parámetros y con ayuda de un programa matemático, calcule la probabilidad de que un posteo tenga más de 140 bytes.

Nota: el número de bytes de un posteo es un valor discreto y, sin embargo, es modelado como una variable aleatoria continua. Esto puede resultar contradictorio, pero es común en la práctica del modelado matemático. Veremos más ejemplos a lo largo de la materia.

32. Desvanecimiento de Rayleigh (*Rayleigh fading*). Al propagarse, la señal de radio se ve afectada por varios factores y, por tanto, la magnitud de la misma al llegar al receptor varía de una forma que puede ser modelada estocásticamente. En un modelo muy simple, la magnitud de la señal recibida R puede modelarse como una variable aleatoria Rayleigh, esto es, $R \sim \text{Rayleigh}(\sigma)$ con $\sigma > 0$, de manera que la función de densidad de probabilidad es

$$f_R(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & x > 0, \\ 0 & x < 0. \end{cases} \tag{31}$$

Encuentre la función de distribución acumulada, el valor esperado y la mediana.

33. Una variable aleatoria X tiene distribución Pareto x_m (> 0) y α (> 0), $X \sim \text{Pareto}(x_m, \alpha)$, si su densidad de probabilidad es

$$\mathrm{f}_X(x) = \begin{cases} \frac{\alpha x_m^\alpha}{x^{\alpha+1}} & x > x_m, \\ 0 & x < x_m. \end{cases} \tag{32}$$

a) Encuentre la función de distribución acumulada y la mediana.

b) Encuentre el valor esperado y la varianza. ¿Para qué valores de α existen?

c) El matemático Vilfredo Pareto notó que, en Italia, el 80 % de la propiedad estaba en manos del 20 % de la población. Esta regla (80-20) es conocida como el principio de Pareto (una búsqueda en Internet puede ser muy iluminadora). Claro que la distribución de Pareto lleva ese nombre debido a la asociación con esa regla. ¿Para qué valor de α vale exactamente la regla si X representa la cantidad de propiedad en manos de una persona tomada al azar?

Ayuda: En vez de propiedades, vamos a pensar en riqueza (por ej., medida en \$). Suponga que hay 100 personas. Ordenemos la riqueza de las personas de menor a mayor: $x_1 \leq x_2 \leq \cdots \leq x_{100}$. Entonces, por ej., la proporción de la riqueza total que posee el 20 % más pobre es

$$L(0.2) = \frac{\sum\limits_{i=1}^{20} x_i}{\sum\limits_{i=1}^{100} x_i} = \frac{\frac{1}{100}\sum\limits_{i=1}^{20} x_i}{\frac{1}{100}\sum\limits_{i=1}^{100} x_i} = \frac{\frac{1}{100}\sum\limits_{i=1}^{20} x_i}{\overline{x}}.$$

Ahora asumamos que, en vez de conocer la riqueza de cada individuo de la población, conocemos cómo se distribuye esa riqueza. Si, por ej. la riqueza tiene un recorrido discreto $\mathcal{R}_X = \{x_1, x_2, \cdots\}$, donde asumimos $x_i \leq x_{i+1}$m tenemos

$$L(0.2) = \frac{\sum\limits_{i=1}^{m_{0.2}} x_i \, \mathrm{P}\left(X = x_i\right)}{\sum\limits_{i=1} x_i \, \mathrm{P}\left(X = x_i\right)},$$

donde

$$m_{0.2} = \text{mín}\left\{ j : \sum_{i=1}^{j} \mathrm{P}\left(X = x_i\right) \geq 0.2 \right\}.$$

Si, por el contrario, la distribución es continua con densidad f_X, tenemos

$$L(0.2) = \frac{\int\limits_0^{x_{0.2}} x \mathrm{f}_X(x) dx}{\int\limits_0^{+\infty} x \mathrm{f}_X(x) dx},$$

donde

$$x_{0.2} = \text{ínf}\left\{ x : \mathrm{F}_X(x) \geq 0.2 \right\}.$$

En general, para una distribución continua, la proporción de la riqueza en manos de la proporción p más pobre está dada por:

$$L(p) = \frac{\int\limits_0^{x_p} x \mathrm{f}_X(x) dx}{\int\limits_0^{+\infty} x \mathrm{f}_X(x) dx},$$

donde

$$x_p = \text{ínf}\left\{ x : \mathrm{F}_X(x) \geq p \right\}.$$

$L(p)$ es conocida como la *curva de Lorenz.* Utilice esta curva para resolver lo pedido.

34. Sea X una variable aleatoria continua con soporte en los reales no negativos, esto es, $\mathrm{F}_X(x) = 0$ para $x \leq 0$. Demuestre que, si $\mathrm{E}\left[X\right]$ existe,

$$\mathrm{E}\left[X\right] = \int\limits_0^{+\infty} (1 - \mathrm{F}_X(x)) dx. \tag{33}$$

35. Sea X una variable aleatoria continua con soporte en los reales no positivos, esto es, $\mathrm{F}_X(x) = 0$ para $x \geq 0$. Demuestre que, si $\mathrm{E}\left[X\right]$ existe,

$$\mathrm{E}\left[X\right] = -\int\limits_{-\infty}^{0} \mathrm{F}_X(x) dx. \tag{34}$$

36. Sea X una variable aleatoria continua. Demuestre que, si $\mathrm{E}[X]$ existe,

$$\mathrm{E}[X] = \int_{0}^{+\infty} (1 - \mathrm{F}_X(x))dx - \int_{-\infty}^{0} \mathrm{F}_X(x)dx = \int_{0}^{+\infty} \mathrm{P}(|X| \leq x)\, dx. \tag{35}$$

5. Respuestas

1 $C = \frac{1}{\pi}$, $\quad \mathrm{E}[X] = 0$, $\quad \mathrm{Var}[X] = \frac{A^2}{2}$, $\quad \mathrm{f}_X(x) = \frac{1}{\pi}(\mathrm{arcsen}(\frac{x}{A}) + \frac{\pi}{2}) \quad -A < x < A$

2. a) $\frac{3}{8}$ b) $\frac{3}{5}$. **3.** $\mathrm{E}[X] = \frac{1}{3}$ $\mathrm{Var}[X] = \frac{1}{18}$, $m = 0.293$.
4. a) $\frac{9}{16}$ b) $\frac{27}{64}$.

5. a) $\mathrm{f}_X(x) = 5\,x^4 - 4\,x^5 \quad 0 < x < 1$ b) 0.342.

6. a) $\frac{1}{4}$ b) $\frac{2}{9}$ c) 2.

7. 0.6.

8. a) 51.29 hs b) 0.6065 y 0.00675 c) 0.2636.

9. 0.039.

10. a) 0.2212 b) 0.9179.

11. $\mathrm{F}_T(t) = 1 - \exp(-n\,t\,/100)$, $\mathrm{f}_T(t) = 0.01\;n\;\exp(-n\,t\,/100)$ $\mathrm{E}[X] = 100/n$.

12. a) 0.001182 1/hora b) 0.00014.

14. a) $\mathrm{f}_X(x) = \lambda\,b\,(\lambda\,x)^{b-1}\,\exp(-(\lambda\,x)^b)$ para $x > 0$ b) 32460 hs. c) 970.

15. e) $Q_1 = \mu - 0.6745\,\sigma, Q_3 = \mu + 0.6745\,\sigma, I_Q = 1.349\,\sigma$. f) $2\,\Phi$(-2.698) ≈ 0.008, aproximadamente 8 de cada 1000 valores sorteados al azar de una variable aleatoria normal se declaran *extremos* con esta regla. g) $\sqrt{\frac{2}{\pi}} \approx 0.798$.

16. a) 97.7 % b) 74.8 % c) 72.5 % d) 20467 e) 16078 f) 0.969 g) 0.2436.

17. a) 2.28 % b) 10.233 $\mu\,F$.

18. 0.964.

19. 19.77 %.

20. $\mu = 1.72$ mm y $\sigma^2 = 0.04$.

21. a) 0.7477 b) 0.0076 .

22. a) 0.55 b) 0.24 c1) 0.02 c2) 0.99.

23. 6.2 años.

24. Aproximadamente 545 hs.

25. 83.55 hs. y 81.77 hs.

26. 7.3 UV.

27. a) 0.8413 b) 85.7 c) 81.5 d) 44 % e) 0.1886.

28. a) 0.335 b) 0.8935.

29. Aproximadamente 0.183.

30. a) $\mathrm{E}[X] = \alpha/\beta$, $\mathrm{Var}[X] = \alpha/\beta^2$ b) 10.3550 mm c) 0.0091.

31. a) mediana$X = e^{\mu}$, b) 225.2025 bytes c) 0.6610.

32. $\mathrm{E}[X] = \sigma\sqrt{pi/2}$, mediana$X = \sigma\sqrt{2\ln(2)}$.

33. a) mediana$(X) = 2^{1/\alpha}x_m$ b) el valor esperado existe para $\alpha > 1$ y la varianza para $\alpha > 2$ c) $\alpha = \ln(0.2)/\ln(0.2/0.8)$.

6. Ejercicios resueltos

Ejercicio 3

La función densidad de probabilidad de la variable aleatoria X es:

$$\mathrm{f}_X(x) = \begin{cases} 2(1-x) & x \in (0,1) \\ 0 & x \notin (0,1) \end{cases}$$

1. Calcular $\mathrm{E}[X]$ y $\mathrm{Var}[X]$.
2. Calcular la mediana m de X definida por $\mathrm{P}(X < m) = 1/2$.

Resolución

El recorrido de la variable aleatoria continua X es el intervalo $(0,1)$. El valor esperado se calcula entonces integrando en ese intervalo el producto $x\,\mathrm{f}_X(x)$. Entonces:

$$\mathrm{E}[X] = \int_0^1 x\,2\,(1-x)\,dx = \frac{1}{3}. \tag{36}$$

La varianza se puede calcular de la siguiente manera: $\mathrm{Var}[X] = \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2$. En este caso

$$\mathrm{E}\left[X^2\right] = \int_0^1 x^2\,2\,(1-x)\,dx = \frac{1}{6}, \tag{37}$$

y la varianza resulta:

$$\mathrm{Var}[X] = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}. \tag{38}$$

La mediana de X está definida por $\mathrm{P}(X < m) = 1/2$. Entonces:

$$\mathrm{P}(X < m) = \int_0^m 2\,(1-x)\,dx = \frac{1}{2}. \tag{39}$$

Resolviendo la integral se obtiene la ecuación $2\,m - m^2 = \frac{1}{2}$ cuya raíz en $(0,1)$ es aproximadamente 0.293.

Ejercicio 6

La duración en horas X de cierto dispositivo electrónico es una variable aleatoria continua con función densidad de probabilidad:

$$\mathrm{f}_X(x) = \begin{cases} \dfrac{100}{x^2} & \text{si } x > 100 \\ 0 & \text{si } x \leq 100 \end{cases} \tag{40}$$

a) ¿Cuál es la probabilidad de que un dispositivo dure menos de 200 hs si se sabe que el dispositivo aún funciona después de 150 hs de servicio?

b) ¿Cuál es la probabilidad de que si se instalan 3 de tales dispositivos en un sistema 2 fallen antes de las 150 horas de servicio y el restante después de las 150 hs?

c) ¿Cuál es el número máximo de dispositivos que se pueden poner en un sistema para que la probabilidad de que después de 150 horas de servicio todos ellos funcionen sea 0.4?

Resolución

Variable aleatoria

La variable aleatoria viene ya definida en el enunciado por:

$$X = \text{Duración (en hs.) de un dispositivo electrónico seleccionado al azar.} \tag{41}$$

Eventos y Datos

No se visualiza ningún otro evento de importancia en el enunciado que aquellos relativos a la distribución de la variable X. Sobre las probabilidades de X, tenemos la función de densidad $\mathrm{f}_X(x)$.

Ítem a

Nos piden:

$$\mathrm{P}\big(X < 200 |\, X > 150\big) = \frac{\mathrm{P}\left(X < 200 \cap X > 150\right)}{\mathrm{P}\left(X > 150\right)} = \frac{\mathrm{P}\left(150 < X < 200\right)}{\mathrm{P}\left(X > 150\right)} \tag{42}$$

Para calcular el denominador, debemos integrar la densidad sobre los valores de x considerados en la región, es decir, para todo $x > 150$:

$$\int_{150}^{+\infty} \mathrm{f}_X(x)\, dx \tag{43}$$

Esta integral es impropia y puede causar dificultades, por lo que apelaremos al recurso del complemento:

$$\mathrm{P}\left(X > 150\right) = 1 - \mathrm{P}\left(X \leq 150\right) = 1 - \mathrm{F}_X(150) = 1 - \int_{-\infty}^{150} \mathrm{f}_X(x)\, dx \tag{44}$$

Como f_X es una función partida, dividimos la integral en dos partes, según los cambios de expresiones:

$$\int_{-\infty}^{150} \mathrm{f}_X(x)\, dx = \int_{-\infty}^{100} \underbrace{\mathrm{f}_X(x)}_{0}\, dx + \int_{100}^{150} \underbrace{\mathrm{f}_X(x)}_{\frac{100}{x^2}}\, dx = \int_{100}^{150} 100x^{-2}\, dx = \left(-100x^{-1}\right)\Big|_{100}^{150} = -\frac{100}{150} - \left(-\frac{100}{100}\right) = \frac{1}{3} \tag{45}$$

Notemos que esta última integral coincide con $\mathrm{F}_X(150)$, entonces podremos utilizar $\mathrm{F}_X(150) = \frac{1}{3}$ como dato más adelante.

Ahora calcularemos el numerador, atendiendo algunas sutilezas, para enfatizar algunos conceptos. Sabiendo que la variable X es continua, sabemos que para todo valor $x \in \mathbb{R}$, $\mathrm{P}(X = x) = 0$. Entonces, denotando M.E. como eventos mutuamente excluyentes:

$$\begin{aligned} \mathrm{P}(150 < X < 200) &= \mathrm{P}(150 < X < 200) + \overbrace{\mathrm{P}(X = 200)}^{0} \\ &\overset{\text{M.E.}}{=} \mathrm{P}(150 < X < 200 \cup X = 200) = \mathrm{P}(150 < X \leq 200). \end{aligned} \tag{46}$$

Si bien no parece tener importancia desde lo operativo, esta sutileza esconde algunos conceptos importantes para comprender algunos objetivos de la materia. Notar que $150 < X < 200$ y $150 < X \leq 200$ son **distintos como eventos** (podemos pensarlos como subconjuntos del espacio muestral). Sin embargo, cuando se calculan las probabilidades de ambos, estos **números son iguales**. Se tiende a decir que “da lo mismo”. Sin embargo, vale aclarar que se refiere a las **probabilidades** y no a los **eventos**.

Por otro lado, también por ser mutuamente excluyentes, se puede obtener la siguiente igualdad:

$$\begin{aligned} \mathrm{P}(X \leq 200) &= \mathrm{P}(X \leq 150 \;\cup\; 150 < X \leq 200) \\ &\overset{\text{M.E.}}{=} \mathrm{P}(X \leq 150) + \mathrm{P}(150 < X \leq 200). \end{aligned} \tag{47}$$

Luego, despejando $\mathrm{P}(150 < X \leq 200)$:

$$\mathrm{P}(150 < X \leq 200) = \mathrm{P}(X \leq 200) - \mathrm{P}(X \leq 150) = \mathrm{F}_X(200) - \mathrm{F}_X(150) \tag{48}$$

Ya sabemos que $\mathrm{F}_X(150) = \frac{1}{3}$, sólo resta determinar $\mathrm{F}_X(200)$:

$$\begin{aligned} \mathrm{F}_X(200) &= \int_{-\infty}^{200} \mathrm{f}_X(x)\,dx = \int_{-\infty}^{100} \underbrace{\mathrm{f}_X(x)}_{0}\,dx + \int_{100}^{200} \underbrace{\mathrm{f}_X(x)}_{\frac{100}{x^2}}\,dx = \int_{100}^{200} 100x^{-2}\,dx \\ &= \left(-100x^{-1}\right)\Big|_{100}^{200} = -\frac{100}{200} - \left(-\frac{100}{100}\right) = \frac{1}{2}. \end{aligned} \tag{49}$$

Por último:

$$\mathrm{P}\big(X < 200 \,|\, X > 150\big) = \frac{\mathrm{F}_X(200) - \mathrm{F}_X(150)}{1 - \mathrm{F}_X(150)} = \frac{\frac{1}{2} - \frac{1}{3}}{1 - \frac{1}{3}} = \frac{\frac{1}{6}}{\frac{2}{3}} = \frac{1}{4} \tag{50}$$

Comentario integrales impropias:

Que en este caso hayamos esquivado la integral impropia no significa que siempre lo podamos hacer. Entonces, resolveremos las siguientes integrales impropias:

$$\begin{aligned} \mathrm{P}(X > 150) &= \int_{150}^{+\infty} \mathrm{f}_X(x)\,dx = \lim_{M \to +\infty} \int_{150}^{M} \frac{100}{x^2}\,dx = \lim_{M \to +\infty} \left(-100x^{-1}\right)\Big|_{150}^{M} \\ &= \lim_{M \to +\infty} \underbrace{-\frac{100}{M}}_{\to 0} - \left(-\frac{100}{150}\right) = \frac{2}{3}. \end{aligned} \tag{51}$$

$$\begin{aligned} \mathrm{P}(X > 200) &= \int_{200}^{+\infty} \mathrm{f}_X(x)\,dx = \lim_{M \to +\infty} \int_{200}^{M} \frac{100}{x^2}\,dx = \lim_{M \to +\infty} \left(-100x^{-1}\right)\Big|_{200}^{M} \\ &= \lim_{M \to +\infty} \underbrace{-\frac{100}{M}}_{\to 0} - \left(-\frac{100}{200}\right) = \frac{1}{2}. \end{aligned} \tag{52}$$

Notar además, que estos cálculos añaden consistencia a nuestros resultados, ya que:

- $\mathrm{P}(X > 150) = 1 - \mathrm{F}_X(150) = 1 - \frac{1}{3} = \frac{2}{3}$
- $\mathrm{P}(X > 200) = 1 - \mathrm{F}_X(200) = 1 - \frac{1}{2} = \frac{1}{2}$

Ítem b

En este ítem, ya que se analiza el funcionamiento de 3 dispositivos, debemos definir una nueva variable aleatoria:

$$Y_3 = \text{Cantidad de dispositivos que fallan antes de las 150 hs en la muestra de 3 dispositivos.} \tag{53}$$

Notar que esta variable es **discreta**, ya que cuenta una cantidad. Al ser una variable discreta, pensemos primero en su recorrido: Puede tener un valor mínimo de 0 y un valor máximo de 3. Es decir, $R_{Y_3} = \{0; 1; 2; 3\}$.

Nos piden $\mathrm{P}(Y_3 = 2)$, por lo que nos vendría bien conocer la distribución de Y_3. Notemos que si consideramos que las 3 extracciones son independientes y la probabilidad de extraer un dispositivo de duración menor a 150 hs como un valor constante p en todas las extracciones, tenemos que $Y_3 \sim \mathrm{Bi}(3, p)$.

En principio, no pareciera que esta distribución discreta Y_3 tuviera alguna relación con la distribución continua X definida en el ejercicio. Sin embargo, notemos que X influye sobre la probabilidad de "éxito":

$$p = \mathrm{P}(X < 150) = \mathrm{P}(X \leq 150) = \mathrm{F}_X(150) = \frac{1}{3} \tag{54}$$

Es decir, $Y_3 \sim \mathrm{Bi}(3; \frac{1}{3})$ y por lo tanto:

$$\mathrm{P}(Y_3 = 2) = \binom{3}{2}\left(\frac{1}{3}\right)^2 \frac{2}{3} = \frac{2}{9} \tag{55}$$

Observación: Una pregunta frecuente es si se puede definir la siguiente variable:

$$Z_3 = \text{Cantidad de dispositivos que fallan } \textbf{después} \text{ de las 150 hs en la muestra de 3 dispositivos.} \tag{56}$$

Esto es completamente válido, sólo que cambia la distribución ($Z_3 \sim \mathrm{Bi}(3, \frac{2}{3})$) el evento buscado ($Z_3 = 1$). Sin embargo, la probabilidad coincide:

$$\mathrm{P}(Z_3 = 1) = \binom{3}{1}\frac{2}{3}\left(\frac{1}{3}\right)^2 = \frac{2}{9} \tag{57}$$

Ítem c

En este ítem, tenemos una variable similar a la anterior, sólo que debemos determinar el número de elementos en la muestra n:

$$Y_n = \text{Cantidad de dispositivos que fallan antes de las 150 hs en la muestra de n dispositivos.} \tag{58}$$

De manera similar a la anterior, podemos asumir que $Y_n \sim \mathrm{Bi}(n, \frac{1}{3})$. Además, nos piden determinar n de forma que **todos** los dispositivos funcionen luego de las 150 hs, por lo que **ninguno** puede fallar antes de las 150 hs. Es decir, nos piden buscar n de forma que $\mathrm{P}(Y_n = 0) = 0.4$.

Vale aclarar lo siguiente, n debe ser **entero**, y esta igualdad puede no cumplirse para valores enteros, por lo que habrá que plantear una desigualdad. En el contexto del problema, la probabilidad deseada representa una situación favorable, ya que se aseguraría que el sistema dura al menos 150 hs. Es decir, nos gustaría que esta probabilidad sea **alta**:

$$
\begin{aligned}
\mathrm{P}\left(Y_n=0\right) \geq 0.4 \Leftrightarrow \underbrace{\binom{n}{0}}_{1}\underbrace{\left(\frac{1}{3}\right)^0}_{1}\left(\frac{2}{3}\right)^n \geq 0.4 \Leftrightarrow \ln\left[\left(\frac{2}{3}\right)^n\right] \geq \ln(0.4) \Leftrightarrow \\
\Leftrightarrow n \cdot \ln\left(\frac{2}{3}\right) \geq \ln(0.4) \overset{\ln(\frac{2}{3})<0}{\Leftrightarrow} n \leq \frac{\ln(0.4)}{\ln(\frac{2}{3})} \approx 2.2598. \qquad (59)
\end{aligned}
$$

Es decir, 2 es el número **máximo** de dispositivos para lograr que el sistema funcione con la probabilidad deseada luego de 150 hs. Tiene sentido que este número sea máximo porque mientras más dispositivos se conecten, más improbable es que funcionen **todos** luego de 150 hs.

Notemos además que $\mathrm{P}\left(Y_2=0\right)=(\frac{2}{3})^2 \approx 0.444>0.4$ y $\mathrm{P}\left(Y_3=0\right)=(\frac{2}{3})^3 \approx 0.296<0.4$, verificando así nuestro resultado.

Ejercicio 10

El tiempo de duración hasta la falla de un dispositivo mecánico se supone que tiene distribución exponencial con una media de 400 hs.

1. Este dispositivo ha funcionado sin fallas durante 400 horas. ¿Cuál es la probabilidad de que falle en las próximas 100 horas?

2. Si se ponen a funcionar 10 de estos dispositivos , ¿cuál es la probabilidad de que falle al menos uno de ellos antes de 100 horas?. Suponga que los dispositivos fallan de manera independiente.

Resolución

La variable aleatoria con distribución exponencial tiene la propiedad de *ausencia de memoria*. Sabiendo que el dispositivo ha funcionado por al menos 400 horas no informa nada respecto de la ocurrencia de fallas en intervalos de tiempo posteriores a 400 hs. La probabilidad de que se produzca una falla en las próximas 100 horas es igual a la probabilidad de que se produzcan fallas en cualquier intervalo de longitud menor o igual que 100 horas. Por consiguiente la probabilidad pedida viene dada por $1 - \exp(-100/400) \approx 0.22120$.

Si se ponen a funcionar 10 de estos dispositivos la probabilidad de que ninguno de ellos falle antes de las 100 horas se puede calcular fácilmente basándose en la independencia de la ocurrencia de fallas. Si se define la variable aleatoria T como el tiempo hasta la falla y la variable aleatoria discreta D como el número de dispositivos que fallan después de las 100 horas de funcionamiento, entonces se tiene:

$$\mathrm{P}\left(D = 10\right) = \left(\mathrm{P}\left(T > 100\right)\right)^{10} = \left(\exp(-100/400)\right)^{10} \approx 0.08209\ . \tag{60}$$

Entonces $\mathrm{P}\left(D < 10\right) \approx 0.91792$.

Ejercicio 12

La duración de ciertos dispositivos es una variable aleatoria con distribución exponencial de parámetro λ (en 1/hora). La probabilidad de que en una muestra de 5 dispositivos haya por lo menos uno que dure más de 1200 hs. es 0.75.

a) Hallar la constante λ.

b) Suponga que estos dispositivos se conectan en serie de manera tal que al fallar por lo menos uno de ellos falla todo el sistema. Calcular la probabilidad de que el sistema funcione más de 1500 horas.

Resolución

Variables aleatorias

Notemos que en este enunciado se definen **dos** variables aleatorias:

- X = duración de un dispositivo seleccionado al azar.
- Y_5 = Cantidad de dispositivos que duran más de 1200 hs en una muestra de 5.

Eventos y Datos

Sabemos que $X \sim \mathcal{E}(\lambda)$ con λ desconocido. Sin embargo, al ser exponencial, sabemos que $\mathrm{F}_X(x) = 1 - e^{-\lambda x}$ si $x > 0$ y $\mathrm{F}_X(x) = 0$ en otro caso.

Notemos que otro dato relevante es $\mathrm{P}(Y_5 \geq 1) = 0.75$. Además, como podemos asumir que el lote es grande y la probabilidad de extraer un dispositivo que dure más de 1200 hs se mantiene constante en cada extracción, se obtiene que $Y_5 \sim \mathrm{Bi}(5; p)$, donde

$$p = \mathrm{P}(X > 1200) = 1 - \mathrm{F}_X(1200) = 1 - (1 - e^{-\lambda \cdot 1200}) = e^{-\lambda \cdot 1200} \tag{61}$$

Ítem a

Considerando el recorrido de Y_5, tenemos que $\mathrm{P}(Y_5 \geq 1) = 1 - \mathrm{P}(Y_5 = 0)$. Es decir,

$$\begin{aligned} \mathrm{P}(Y_5 \geq 1) = 1 - \mathrm{P}(Y_5 = 0) = 0.75 \Rightarrow \mathrm{P}(Y_5 = 0) = 0.25 & \\ \Rightarrow \underbrace{\binom{5}{0}}_{1} \underbrace{\left(e^{-\lambda \cdot 1200}\right)^0}_{1} \left(1 - e^{-\lambda \cdot 1200}\right)^5 = 0.25 & \end{aligned} \tag{62}$$

A partir de esta expresión podemos despejar λ:

$$\begin{aligned} \left(1 - e^{-\lambda \cdot 1200}\right)^5 = 0.25 \Rightarrow 1 - e^{-\lambda \cdot 1200} = \sqrt[5]{0.25} \Rightarrow e^{-\lambda \cdot 1200} = 1 - \sqrt[5]{0.25} & \\ \Rightarrow -\lambda \cdot 1200 = \ln(1 - \sqrt[5]{0.25}) & \end{aligned} \tag{63}$$

Por último,

$$\lambda = -\frac{\ln(1 - \sqrt[5]{0.25})}{1200} \approx 0.0011819 \tag{64}$$

Ítem b

Ahora que sabemos λ, tenemos caracterizada la distribución de X. Ahora, tenemos una variable similar a la anterior:

$$Z_5 = \text{ Cantidad de dispositivos que duran más de } \underline{1500 \text{ hs}} \text{ en una muestra de 5.} \tag{65}$$

La única diferencia respecto a Y_5 es el tiempo de duración. Con argumentos similares, $Z_5 \sim \mathrm{Bi}(5, p)$, donde

$$p = \mathrm{P}\left(X > 1500\right) = e^{-0.0011819 \cdot 1500} \tag{66}$$

Entonces, nos piden la probabilidad de que todos los dispositivos funcionen luego de 1500 hs. Es decir,

$$\mathrm{P}\left(Z_5 = 5\right) = \underbrace{\binom{5}{5}}_{1} \left(e^{-0.0011819 \cdot 1500}\right)^5 \underbrace{\left(1 - e^{-0.0011819 \cdot 1500}\right)^0}_{1} \approx 0.00014135 \tag{67}$$

Ejercicio 14

Se sabe que la duración X de una pieza responde a la distribución de Weibull cuya función de distribución es: $\mathrm{F}_X(x) = 1 - \exp(-(\lambda x)^b)$ para $x > 0$ con b y λ constantes reales positivas. Si x se mide en miles de horas, $b = 2$ y $\lambda = 0.01$:

a) obtener la función de densidad de probabilidad de la duración de la pieza.

b) calcular la duración garantizada G_{90} con un 90 % de confiabilidad. Nota: Se cumple $P(X > G_{90}) = 0.90$.

c) Se realizó una prueba con un lote muy grande de piezas y se separaron 1000 con duración superior a 10000 horas; ¿cuántas unidades espera Ud. haya en este lote de 1000 cuyas duraciones totales sean superiores a 20000 horas?

Resolución

Variable aleatoria

La variable aleatoria viene definida en el enunciado:

$$X = \text{ Duración de una pieza elegida al azar en miles de hs.} \tag{68}$$

El hecho de que las unidades estén en miles de horas, nos dice que si para una pieza $X > 10$, entonces esa pieza dura más de 10000 hs.

Datos

Tenemos la distribución acumulada de la variable X. Como se dice que $b = 2$ y $\lambda = 0.01$, $\mathrm{F}_X(x) = 1 - e^{-(0.01x)^2}$ si $x > 0$ y $\mathrm{F}_X(x) = 0$ si $x \leq 0$.

Ítem a

Para obtener la densidad, podemos derivar la acumulada, por lo tanto, si $x > 0$:

$$\begin{aligned} \mathrm{f}_X(x) &= [\mathrm{F}_X(x)]' = [1 - e^{-(0.01x)^2}]' = -e^{-(0.01x)^2} \cdot [-(0.01x)^2]' \\ &= -e^{-(0.01x)^2} \cdot (-2 \cdot 0.01x \cdot [0.01x]') = 0.0002 \cdot x \cdot e^{-(0.01x)^2} \end{aligned} \tag{69}$$

Por otro lado, como $\mathrm{F}_X(x) = 0$ si $x \leq 0$, entonces la derivada también es nula:

$$\mathrm{f}_X(x) = \begin{cases} 0.0002 \cdot x \cdot e^{-(0.01x)^2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \tag{70}$$

Ítem b

Nos piden hallar un valor G_{90} de forma que $\mathrm{P}(X > G_{90}) = 0.90$. Para eso, utilizaremos la función de distribución:

$$\mathrm{P}(X > G_{90}) = 0.9 \Rightarrow 1 - \mathrm{F}_X(G_{90}) = 0.9 \Rightarrow 1 - (1 - e^{-(0.01G_{90})^2}) = 0.9 \Rightarrow e^{-(0.01G_{90})^2} = 0.9 \tag{71}$$

Por lo tanto, para determinar G_{90}, utilizamos el logaritmo:

$$\begin{aligned} e^{-(0.01G_{90})^2} = 0.9 &\Rightarrow -(0.01G_{90})^2 = \ln(0.9) \Rightarrow 0.01G_{90} = \sqrt{-\ln(0.9)} \\ &\Rightarrow G_{90} = \frac{\sqrt{-\ln(0.9)}}{0.01} \approx 32.459 \end{aligned} \tag{72}$$

Es decir, como G_{90} es 32.459 y está medido en miles de horas, se puede garantizar el funcionamiento de una pieza durante 32459 hs con una confianza del 90 %.

Ítem c

Aquí sabemos que las 1000 piezas seleccionadas tienen una duración mayor a 10000 hs (es decir, $X > 10$). Queremos estimar cuántas de estas cumplen $X > 20$. Podríamos pensar en las siguiente variable:

$$Y_{1000} = \text{Cantidad de piezas con duración mayor a 20000 hs entre las 1000 seleccionadas} \tag{73}$$

Podríamos pensar esta variable como una binomial, ya que se extraen las piezas de un lote muy grande y la probabilidad de extraer una pieza con una duración mayor a 20000 hs se mantiene con valores similares para las 1000 extracciones.

Sin embargo, cada una de estas sabemos que tienen una duración mayor a 10000 hs. Por lo que la probabilidad de "éxito" de esta binomial será **condicional**. Es decir, $Y_{1000} \sim \text{Bi}(1000, p)$ donde $p = \text{P}\big(X > 20 | X > 10\big)$.

Calcularemos ahora dicha probabilidad p :

$$\begin{aligned} \text{P}\big(X > 20 | X > 10\big) &= \frac{\text{P}\left(X > 20 \cap X > 10\right)}{\text{P}\left(X > 10\right)} = \frac{\text{P}\left(X > 20\right)}{\text{P}\left(X > 10\right)} = \frac{1 - \text{F}_X(20)}{1 - \text{F}_X(10)} \\ &= \frac{e^{-(0.01 \cdot 20)^2}}{e^{-(0.01 \cdot 10)^2}} = \frac{e^{-0.04}}{e^{-0.01}} \\ &\approx 0.97045. \end{aligned} \tag{74}$$

Entonces, el número esperado de piezas con duración mayor a 20000 hs será:

$$\text{E}\left[Y_{1000}\right] = 1000 \cdot 0.97045 = 970.45 \tag{75}$$

Como debe ser un número entero, redondearemos al más cercano y esperamos un número aproximado de 970 piezas con duración mayor a 20000 hs.

Ejercicio 21

El diámetro de contacto de la rosca de una unión se distribuye normalmente con media 10.02 mm y desvío estándar 0.075 mm. Las especificaciones dadas para esa rosca son: 10.00 $\pm$ 0.025 mm. (la especificación es un intervalo de valores dentro del cual deberá estar el diámetro para que sea considerado aceptable).

a) ¿Cuál es la probabilidad de que la rosca esté fuera de la especificación dada?

b) Suponiendo que el diámetro de contacto esté centrado en el valor nominal de la especificación (10 mm), ¿cuál es la máxima desviación estándar del diámetro aceptable que permitirá no más de una defectuosa entre mil producidas?

Resolución

Variable aleatoria

La variable aleatoria es:

$$D = \text{ Diámetro de contacto de una rosca seleccionada al azar.} \tag{76}$$

Eventos

Podemos considerar el evento:

$$E = \text{La rosca está dentro de los límites de especificación.} \tag{77}$$

Sin embargo, este evento lo podemos vincular con la variable D, ya que $E = 9.975 \leq D \leq 10.025$.

Datos

El dato que sabemos es que D tiene una distribución normal, cuyos parámetros también son dados: $D \sim N(10.02; 0.075)$.

Ítem a

Nos piden la probabilidad de que una rosca esté fuera de los límites de especificación. Es decir $\mathrm{P}\left(\overline{E}\right) = 1 - \mathrm{P}\left(E\right)$. Por lo tanto, podemos calcular $\mathrm{P}\left(E\right)$ primero:

$$\mathrm{P}\left(E\right) = \mathrm{P}\left(9.975 \leq D \leq 10.025\right) \tag{78}$$

Como D tiene una distribución normal, sus probabilidades se pueden calcular en base a la distribución normal estándard, restando la media (10.02) y dividiendo por el desvío (0.075):

$$\mathrm{P}\left(E\right) = \mathrm{P}\left(9.975 \leq D \leq 10.025\right) = \mathrm{P}\left(\frac{9.975 - 10.02}{0.075} \leq \frac{D - 10.02}{0.075} \leq \frac{10.025 - 10.02}{0.075}\right) \tag{79}$$

Esta última igualdad entre probabilidades, se debe a que como eventos deben ser iguales, es decir, siempre que se cumpla $9.975 \leq D \leq 10.025$ se cumple también $\frac{9.975-10.02}{0.075} \leq \frac{D-10.02}{0.075} \leq \frac{10.025-10.02}{0.075}$ y viceversa, ya que se aplican funciones crecientes y biyectivas en todos los miembros de la desigualdad. Si bien luego es algo que se incorpora, es importante saber los motivos por los que las probabilidades no se ven alteradas, ya que aplicar transformaciones erróneas puede derivar en un cálculo desacertado de probabilidades.

Por otro lado, como $D \sim N(10.02; 0.075) \Rightarrow \frac{D-10.02}{0.075} \sim N(0;1)$. Entonces:

$$\begin{aligned}
&\mathrm{P}\left(\underbrace{\frac{9.975-10.02}{0.075}}_{-0.6} \leq \frac{D-10.02}{0.075} \leq \underbrace{\frac{10.025-10.02}{0.075}}_{0.066}\right) = \\
=&\mathrm{P}\left(\frac{D-10.02}{0.075} \leq 0.066\right) - \mathrm{P}\left(\frac{D-10.02}{0.075} \leq -0.6\right) = \\
=&\Phi(0.066) - \Phi(-0.6). \qquad (80)
\end{aligned}$$

De estos últimos valores, la primera se puede calcular por tabla y se corresponde con la siguiente área:

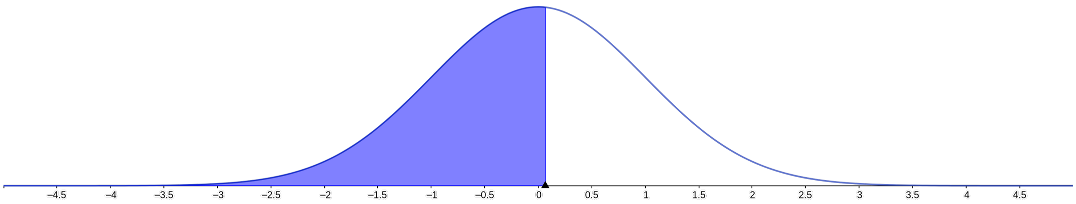

Figura 1: Área correspondiente a $\Phi(0.066) = 0.5239$

Para calcular $\Phi(-0.6)$, como la tabla no tiene valores negativos, nos valdremos de la simetría de la distribución normal:

Por lo tanto,

$$\mathrm{P}(E) = \Phi(0.066) - \Phi(-0.6) = 0.5239 - 0.2743 = 0.2536 \Rightarrow \mathrm{P}\left(\overline{E}\right) = 1 - 0.2536 = 0.7464 \qquad (81)$$

Ítem b

En este nuevo ítem, la media es conocida pero el desvío no. Por lo tanto, $D \sim N(10; \sigma)$.

Se pide σ de forma que $\mathrm{P}\left(\overline{E}\right) = 0.001$. Esto equivale a pedir que $\mathrm{P}(E) = 0.999$:

$$0.999 = \mathrm{P}(E) = \mathrm{P}(9.975 \leq D \leq 10.025) = \mathrm{P}\left(\frac{9.975-10}{\sigma} \leq \frac{D-10}{\sigma} \leq \frac{10.025-10}{\sigma}\right) \qquad (82)$$

Notar que hemos restado y dividido por otros valores ya que en este ítem, la media es de 10 mm y el desvío desconocido. Con estos últimos cálculos, hemos estandarizado la variable D y por lo tanto:

$$\begin{aligned}
0.999 = \mathrm{P}(E) &= \Phi\left(\frac{0.025}{\sigma}\right) - \Phi\left(-\frac{0.025}{\sigma}\right) \overset{\text{Simetría}}{=} \Phi\left(\frac{0.025}{\sigma}\right) - \left[1 - \Phi\left(\frac{0.025}{\sigma}\right)\right] \\
&= 2\Phi\left(\frac{0.025}{\sigma}\right) - 1 \qquad (83)
\end{aligned}$$

Para despejar σ, utilizaremos la última expresión:

$$0.999 = 2\Phi\left(\frac{0.025}{\sigma}\right) - 1 \Rightarrow \frac{0.999+1}{2} = \Phi\left(\frac{0.025}{\sigma}\right) \Rightarrow \Phi^{-1}(0.9995) = \frac{0.025}{\sigma} \Rightarrow \qquad (84)$$

$$\Rightarrow \sigma = \frac{0.025}{\Phi^{-1}(0.9995)} \approx \frac{0.025}{3.0902} \approx 0.008090091 \qquad (85)$$

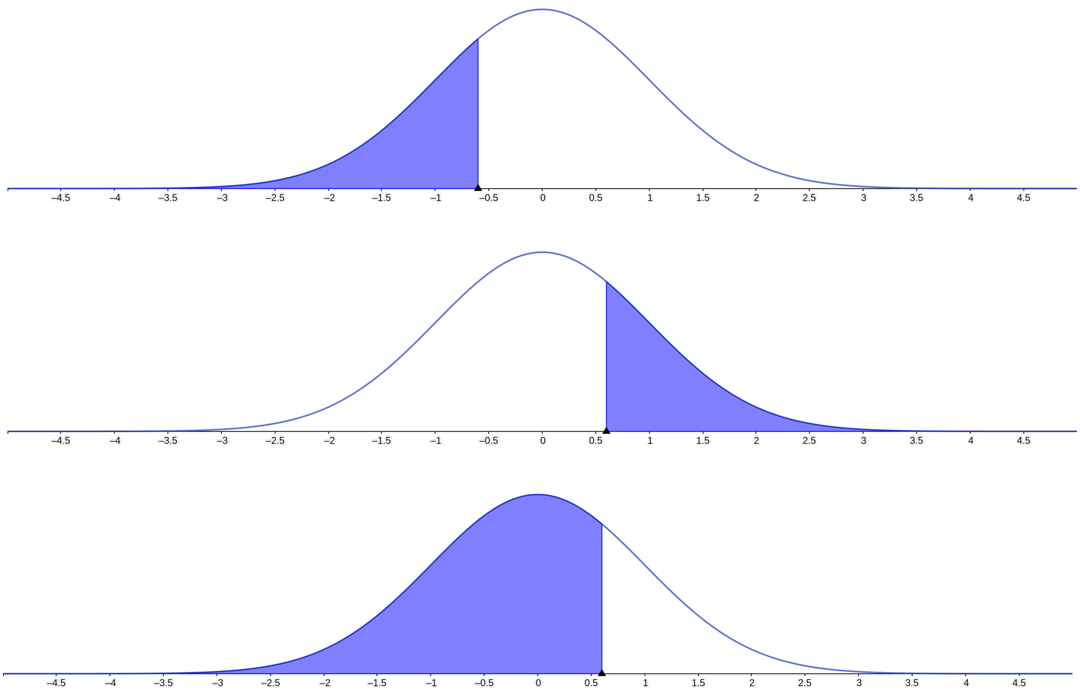

Figura 2: El primer gráfico corresponde al área que equivale a $\Phi(-0.6)$, que coincide con el área del segundo gráfico. Este valor se obtiene como el complemento de $\Phi(0.6)$, representado en el tercer gráfico. Es decir, se concluye que $\Phi(-0.6) = 1 - \Phi(0.6) = 1 - 0.7257 = 0.2743$.

Observación: El valor del cuantil fue obtenido con el último valor de la tabla, porque era el mayor disponible. Pero si se utiliza un software, este valor puede calcularse de forma aún más precisa:

$$\Rightarrow \sigma = \frac{0.025}{\Phi^{-1}(0.9995)} \approx \frac{0.025}{3.2905} \approx 0.0075976 \tag{86}$$

Ejercicio 27

Se supone que los resultados de un examen tienen una distribución normal con una media de 78 y una varianza de 36.

1. Determinar la probabilidad de que una persona que realiza el examen obtenga una nota mayor que 72.
2. Suponga que a los estudiantes que se encuentran en el 10 % de la parte superior de la distribución se les asigna una calificación A. ¿Cuál es la nota mínima que se debe obtener para tener calificación A?
3. ¿Cuál debe ser la mínima nota aprobatoria si el evaluador pretende que solamente el 28.1 % de los estudiantes apruebe?
4. Calcular, aproximadamente, la proporción de los estudiantes que tienen nota que exceden por lo menos en 5 puntos a la calificación reprobatoria del 25 % (calificaciones inferiores).
5. Si se sabe que la calificación de un estudiante es mayor que 72, ¿cuál es la probabilidad de que su calificación sea mayor que 84?

Resolución

1. Llamemos

$$X = \text{calificación en el examen de una persona tomada al azar.} \tag{87}$$

Según el enunciado, $X \sim \mathcal{N}(\mu_X, \sigma_X)$, donde $\mu_X = 78$ y $\sigma_X = \sqrt{36} = 6$. Entonces,

$$\begin{aligned} \mathrm{P}(X > 72) &= 1 - \mathrm{P}(X \leq 72) = 1 - \mathrm{P}\left(\frac{X-78}{6} \leq \frac{72-78}{6}\right) = 1 - \Phi(-1) = \Phi(1) \\ &= 0.8413, \end{aligned} \tag{88}$$

donde el valor de $\Phi(1)$ puede ser obtenido de una tabla de distribución normal o de una calculadora. Usando, por ejemplo, el programa Octave, podemos hacer

```
>> normcdf(1)
ans =  0.84134
```

También podríamos haber hecho

```
>> 1-normcdf(-1)
ans =  0.84134
```

O, incluso, sin normalizar,

```
>> 1-normcdf(72,78,6)
ans =  0.84134
```

En este caso, es segundo parámetro corresponde a la media y el tercero al desvío.

2. Otra forma equivalente de plantear la pregunta es la siguiente: ¿cuál es la calificación mínima que es superada por un alumno tomado al azar con una probabilidad de 0.1? Es decir, buscamos x_A tal que

$$\mathrm{P}(X \geq x_A) = 0.1. \tag{89}$$

Primero debemos observar que, dado que X es una variable aleatoria continua,

$$\mathrm{P}\left(X \geq x_A\right) = \mathrm{P}\left(X > x_A\right) = 1 - \mathrm{P}\left(X \leq x_A\right). \tag{90}$$

Luego,

$$\mathrm{P}\left(X \geq x_A\right) = 1 - \Phi\left(\frac{x_A - 78}{6}\right) = 0.1 \Rightarrow \Phi\left(\frac{x_A - 78}{6}\right) = 0.9 \Rightarrow \tag{91}$$

$$\frac{x_A - 78}{6} = \Phi^{-1}(0.9) = z_{0.9} \Rightarrow x_A = 78 + 6z_{0.9}. \tag{92}$$

Buscando en una tabla de la inversa de la distribución Gaussiana o usando la calculadora, encontramos que $z_{0.9} \approx 1.2816$. Luego,

$$x_A = 85.6893 \approx 85.7. \tag{93}$$

Otra forma de encontrar el valor de la inversa de la distribución normal es usando un programa de computadora como, por ej., Octave:

```
>> norminv(0.9)
ans =  1.2816
```

También se podría haber hecho, directamente,

```
>> norminv(0.9,78,6)
ans =  85.689
```

En este caso, hemos pasado como argumentos la media y el desvío.

3. Llamemos x_P a la calificación que permitiría pasar a los alumnos. Si el docente desea que sólo el 28.1 % de los alumnos apruebe, x_P debería ser tal

$$\mathrm{P}\left(X \geq x_P\right) = 0.281. \tag{94}$$

Luego,

$$\mathrm{P}\left(X \geq x_P\right) = 1 - \mathrm{P}\left(X \leq x_P\right) = 1 - \Phi\left(\frac{x_P - 78}{6}\right) = 0.281 \Rightarrow \Phi\left(\frac{x_P - 78}{6}\right) = 0.719. \tag{95}$$

$$\Rightarrow \frac{x_P - 78}{6} = 0.719 \Rightarrow x_P = 78 + 6z_{0.719} \approx 78 + 6 \cdot 0.5799 = 81.4794 \approx 81.5, \tag{96}$$

donde hemos buscado en una tabla el valor $z_{0.719} \approx 0.5799$.

4. Si llamamos x_R a la calificación reprobatoria, tenemos

$$\mathrm{P}\left(X \leq x_R\right) = \Phi\left(\frac{x_R - 78}{6}\right) = 0.25 \Rightarrow x_R = 78 + 6z_{0.25}. \tag{97}$$

En muchas tablas, no aparece $\Phi^{-1}(\alpha)$ para $\alpha < 0.5$. El hecho es que la simetría de la densidad de probabilidad de una variable aleatoria normal permite utilizar una tabla más pequeña que sólo muestre los resultados para valores de $\alpha \geq 0.5$. De hecho, sabemos que

$$\Phi^{-1}(0.25) = -\Phi^{-1}(1 - 0.25) = -\Phi^{-1}(0.75). \tag{98}$$

Entonces,

$$x_R = 78 - 6z_{0.75} \approx 78 - 6 \cdot 0.6745 = 73.9530, \tag{99}$$

donde hemos obtenido en la tabla que $z_{0.75} \approx 0.6745$.

Finalmente, podemos responder la pregunta del ejercicio:

$$\mathrm{P}\left(X \geq x_R + 5\right) = \mathrm{P}\left(X \geq 78.9530\right) = 1 - \Phi\left(\frac{78.9530 - 78}{6}\right) \approx 1 - 0.5631 = 0.4369. \tag{100}$$

Es decir, aproximadamente el 44 % de los estudiantes tiene una calificación que supera en por lo menos 5 puntos a la nota reprobatoria del 25 %.

5. La construcción

   “Si tal condición se da, ¿cuál es la probabilidad de tal evento?”

   claramente nos indica que se trata de una probabilidad condicional. Es decir, el ejercicio pide la siguiente probabilidad

$$\begin{aligned}
\mathrm{P}\big(\, X > 84 |\, X > 72\big) &= \frac{\mathrm{P}\left(\{X > 84\} \cap \{X > 72\}\right)}{\mathrm{P}\left(X > 72\right)} = \frac{\mathrm{P}\left(X > 84\right)}{\mathrm{P}\left(X > 72\right)} \\
&= \frac{1 - \Phi\left(\frac{84-78}{6}\right)}{1 - \Phi\left(\frac{72-78}{6}\right)} = \frac{1 - \Phi\left(+1\right)}{1 - \Phi\left(-1\right)} = \frac{1 - \Phi\left(+1\right)}{\Phi\left(+1\right)} \\
&= \frac{1 - 0.8413}{0.8413} \approx 0.1886,
\end{aligned} \tag{101}$$

   donde hemos usado la simetría de la distribución normal y buscado en la tabla el valor $\Phi(1) \approx 0.8413$.

Ejercicio 29

La longitud de ciertos dispositivos mecánicos tiene la siguiente tolerancia: 58 ± 0.21 mm. Se han recibido tres partidas de distintos proveedores de 200, 300 y 500 cajas de 6 unidades cada una, que si bien no difieren en su valor medio 58 mm presentan distintas dispersiones: 0.102 mm, 0.12 mm y 0.146 mm. Se supone que la longitud de esos dispositivos es una variable aleatoria con distribución normal. En el almacén se han mezclado las cajas. Si se elige una caja al azar y se comprueba que exactamente 4 de las piezas están dentro de los límites de tolerancia, ¿cuál es la probabilidad de que la caja haya sido entregada por el proveedor número 2?

Resolución

Variables aleatorias

Podemos considerar las siguientes variables aleatorias:

- X = la longitud de un dispositivo seleccionado al azar.
- X_i = la longitud de un dispositivo del proveedor i seleccionado al azar. $(1 \le i \le 3)$
- Y = Cantidad de dispositivos con longitud dentro de los límites de tolerancia entre los 6 de la caja seleccionada.
- Y_i = Cantidad de dispositivos con longitud dentro de los límites de tolerancia entre los 6 de una caja seleccionada del proveedor i. $(1 \le i \le 3)$

Eventos

Definimos los siguientes eventos:

- A_i = la caja seleccionada viene del proveedor i. $(1 \le i \le 3)$
- T_i = Un dispositivo seleccionado al azar del proveedor i tiene una longitud dentro de los límites de tolerancia. $(1 \le i \le 3)$

Datos

Tenemos los siguientes datos:

- $\mathrm{P}(A_1) = 0.2$
- $\mathrm{P}(A_2) = 0.3$
- $\mathrm{P}(A_3) = 0.5$
- $X_1 \sim N(58; 0.102)$
- $X_2 \sim N(58; 0.12)$
- $X_3 \sim N(58; 0.146)$

Además, podemos considerar que cada variable Y_i tiene una distribución binomial, con probabilidad de "éxito" $p_i = \mathrm{P}(T_i)$. Es decir, $Y_i \sim \mathrm{Bi}(6; p_i)$.

Nos piden $\mathrm{P}(A_2 | Y = 4)$. Utilizando el teorema de Bayes:

$$\mathrm{P}(A_2 | Y = 4) = \frac{\mathrm{P}(Y = 4 | A_2)\mathrm{P}(A_2)}{\mathrm{P}(Y = 4 | A_1)\mathrm{P}(A_1) + \mathrm{P}(Y = 4 | A_2)\mathrm{P}(A_2) + \mathrm{P}(Y = 4 | A_3)\mathrm{P}(A_3)} \tag{102}$$

Es decir,

$$\mathrm{P}(A_2 | Y = 4) = \frac{\mathrm{P}(Y_2 = 4)\mathrm{P}(A_2)}{\mathrm{P}(Y_1 = 4)\mathrm{P}(A_1) + \mathrm{P}(Y_2 = 4)\mathrm{P}(A_2) + \mathrm{P}(Y_3 = 4)\mathrm{P}(A_3)} \tag{103}$$

Es decir, nos resta calcular $\mathrm{P}(Y_i = 4)$ para $1 \le i \le 3$. Sabiendo que son binomiales, resta sacar la probabilidad $\mathrm{P}(T_i)$.

- **Si** $i = 1$:

$$\begin{aligned} \mathrm{P}(T_1) &= \mathrm{P}(58 - 0.21 \leq X_1 \leq 58 + 0.21) = \Phi\left(\frac{0.21}{0.102}\right) - \Phi\left(-\frac{0.21}{0.102}\right) \\ &= 2\Phi\left(\frac{0.21}{0.102}\right) - 1 \approx 0.96049 \end{aligned} \tag{104}$$

$$\mathrm{P}(Y_1 = 4) = \binom{6}{4}(0.96049)^4(1 - 0.96049)^2 \approx 0.019929 \tag{105}$$

- **Si** $i = 2$:

$$\begin{aligned} \mathrm{P}(T_2) &= \mathrm{P}(58 - 0.21 \leq X_2 \leq 58 + 0.21) = \Phi\left(\frac{0.21}{0.12}\right) - \Phi\left(-\frac{0.21}{0.12}\right) \\ &= 2\Phi\left(\frac{0.21}{0.12}\right) - 1 \approx 0.91988 \end{aligned} \tag{106}$$

$$\mathrm{P}(Y_2 = 4) = \binom{6}{4}(0.91988)^4(1 - 0.91988)^2 \approx 0.068944 \tag{107}$$

- **Si** $i = 3$:

$$\begin{aligned} \mathrm{P}(T_3) &= \mathrm{P}(58 - 0.21 \leq X_3 \leq 58 + 0.21) = \Phi\left(\frac{0.21}{0.146}\right) - \Phi\left(-\frac{0.21}{0.146}\right) \\ &= 2\Phi\left(\frac{0.21}{0.146}\right) - 1 \approx 0.84967 \end{aligned} \tag{108}$$

$$\mathrm{P}(Y_3 = 4) = \binom{6}{4}(0.84967)^4(1 - 0.84967)^2 \approx 0.17668 \tag{109}$$

Notar que a mayor dispersión, menor es la probabilidad de que un dispositivo esté dentro de los límites de tolerancia.

Por lo tanto,

$$\mathrm{P}\big(A_2 | Y = 4\big) \approx \frac{0.068944 \cdot 0.3}{0.019929 \cdot 0.2 + 0.068944 \cdot 0.3 + 0.17668 \cdot 0.5} \approx 0.18302 \tag{110}$$