Probabilidad y estadística

MAT 041, Primer semestre

Francisco Cuevas Pacheco

27 de septiembre de 2022

Contenidos

Distribución Gamma

Sean α y β parámetros positivos. Una variable aleatoria **gamma** tiene función de densidad dada por

$$f_X(x) = \frac{\beta(\beta x)^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, \text{ si } x > 0$$

Usamos la notación $X \sim \text{Gamma}(\alpha, \beta)$. Aquí, $\Gamma(\cdot)$ es la función Gamma:

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du,$$

la cual satisface las siguientes propiedades

- $\Gamma(n) = (n-1)!$ para $n \in \mathbb{N}$
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

Distribución Gamma

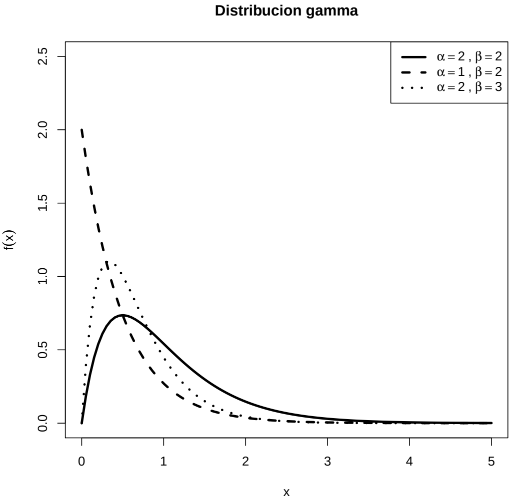


Figura 1: Función de densidad de $X \sim \text{Gamma}(\alpha, \beta)$ para diferentes α y β.

Distribución Gamma

La esperanza de X se obtiene desde las siguientes igualdades

$$\begin{aligned} E(X) &= \int_0^\infty x \cdot \frac{\beta(\beta x)^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\beta x} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \end{aligned}$$

Por lo tanto,

$$\boxed{E(X) = \frac{\alpha}{\beta}}$$

La tercera igualdad es una identidad que generaliza la siguiente transformada de Laplace:

$$\int_0^\infty x^n e^{-sx} dx = \frac{n!}{s^{n+1}}, \qquad n \in \mathbb{N}.$$

Distribución Gamma

Similarmente,

$$\begin{aligned} E(X^2) &= \int_0^\infty x^2 \cdot \frac{\beta(\beta x)^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1} e^{-\beta x} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{(\beta)^{\alpha+2}} \\ &= (\alpha+1)\alpha\beta^2 \end{aligned}$$

Por lo tanto, la varianza está dada por

$$\boxed{\operatorname{var}(X) = \frac{\alpha}{\beta^2}}$$

Relación Gamma - Exponencial

Recordemos que los tiempos de ocurrencia de un proceso Poisson, T_i, tienen una distribución exponencial de parámetro λ. Luego se cumple que

$$S_n = T_1 + \ldots + T_n \sim \text{Gamma}(n, \lambda)$$

De aquí se desprende que $S_1 \sim \text{Exp}(\lambda) = \text{Gamma}(1, \lambda)$.

Ejercicio

En astronomía, una bola de fuego es un meteoríto con un brillo inusual que posee más de 0.1 *kilotones de energía. Se sabe que el número de bolas de fuego detectadas por los sistemas satelitales es bien modelado por un proceso Poisson con una tasa de 2 bolas de fuego por año.*

(c) Calcule la probabilidad de que el tiempo para observar 3 bolas de fuego no supere un año.

(d) ¿Cuál es la probabilidad de observar 6 bolas de fuego en a lo más 3 años?

(e) ¿Cuántos años se deben esperar para observar 10 bolas de fuego?

Solución:

Definimos a $S_n = T_1 + \ldots + T_n$. Luego $S_n \sim \text{Gamma}(n, 2)$

(c) $P(S_3 \leq 1) = 0.32332$

(d) $P(S_6 \leq 3) = 0.55432$

(e) $E(S_{10}) = 5$

Distribución Normal

Una variable aleatoria X sigue una distribución **normal** si su función de densidad está dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R},$$

donde $\mu \in \mathbb{R}$ y $\sigma > 0$ son parámetros del modelo. Usaremos la notación $X \sim \mathrm{N}(\mu, \sigma^2)$.

Para la distribución normal tenemos que la esperanza y la varianza están dadas por $E(X) = \mu$ y $\mathrm{var}(X) = \sigma^2$, respectivamente.

Distribución Normal

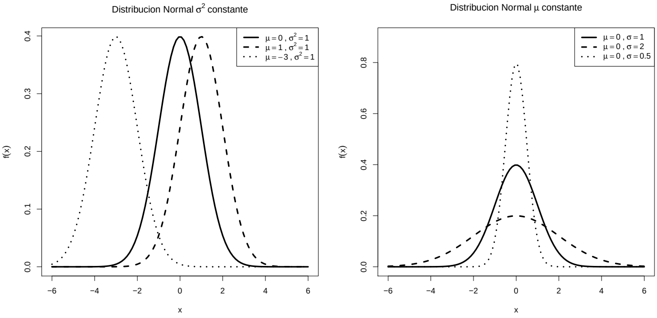


Figura 2: Función de densidad de $X \sim \mathrm{N}(\mu, \sigma^2)$ para diferentes valores de μ y σ^2.

Observaciones

- Cuando $\mu = 0$ y $\sigma^2 = 1$, la variable aleatoria X se llama **normal estándar**, en cuyo caso, la función de densidad es simplemente

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}.$$

- Sea $X \sim \mathrm{N}(\mu, \sigma^2)$. El teorema de cambio de variable implica que

$$Z = \frac{X - \mu}{\sigma} \sim \mathrm{N}(0, 1).$$

Podemos calcular las probabilidades de la distribución normal a partir de las probabilidades de la distribución normal estándar. Por ejemplo, sea $X \sim \mathrm{N}(\mu, \sigma^2)$, entonces

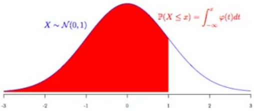
$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right)$$

Si usamos la notación

$$\Phi(z) = \int_{-\infty}^{z} \phi(x)\mathrm{d}x,$$

para la función de distribución de la normal estándar (esta función está tabulada), la probabilidad anterior se puede calcular de la siguiente forma

$$P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$



	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

Ejemplo

Sea $X \sim N(0, \sigma^2)$. Calcule la probabilidad de que

(a) $|X| \leq \sigma$

(b) $X \leq 1.65\sigma$

(c) $|X| \leq 2\sigma$

(d) $|X| \leq 3\sigma$

(e) $X > 1.13\sigma$

(f) $X > 1.108\sigma$

Ejemplo

Solución:

Se sabe que $Z = X/\sigma \sim N(0, 1)$. Luego

(a) $P(|X| \leq \sigma) = P(-1 \leq Z \leq 1) = 0.8413 - 0.1587 = 0.6826$

(b) $P(X \leq 1.65\sigma) = P(Z \leq 1.65) = 0.9505$

(c) $P(|X| \leq 2\sigma) = P(-2 \leq Z \leq 2) = 0.9772 - 0.0228 = 0.9544$

(d) $P(|X| \leq 3\sigma) = P(-3 \leq Z \leq 3) = 0.9987 - 0.0013 = 0.9974$

(e) $P(X > 1.13) = 1 - P(X \leq 1.13) = 1 - 0.8708 = 0.1292$

(f) $P(X > 1.108) = 1 - P(X \leq 1.108) = 1 - \frac{0.8643+0.8665}{2} = 1 - 0.8654 = 0.1346$

Ejercicio

Se sabe que la distribución de la estatura de un hombre chileno es aproximadamente Normal, de media $1.65[M]$ *y desviación estandar* $0.1[M]$*. ¿Qué proporción de personas tiene una estatura mayor a* $1.95[M]$*?*

Solución:

Sea X: estatura de un hombre chileno. $X \sim N(\mu, \sigma^2)$ y en consecuencia $(Z - \mu)/\sigma \sim N(0,1)$. Luego

$$P(X \leq 1.95) = P\left(\frac{X - 1.65}{0.1} \leq \frac{1.95 - 1.65}{0.1}\right) = P\left(Z \leq 3\right) = 0.9987$$

Ejercicio

Se sabe que la resistencia a la ruptura de un cierto tipo de cable, X medida en [MP], sigue una distribución Normal. Se ha determinado que un 84.13 % de los cables poseen resistencia menor que 114 [MP] y que sólo un 2.28 % de ellos tiene una resistencia de, a lo menos, de 118 [MP]. Se pide:

a) *Determinar la resistencia esperada y la varianza de la resistencia del cable.*

b) *La utilidad obtenida por la venta de un cable es el 15 % del precio de venta. Determine la utilidad esperada y su función de distribución si el precio de venta del cable Y, en [K\$], está dado por:*

$$Y(X) = \begin{cases} 1000 & X \leq 105 \\ 3000 & 105 < X < 120 \\ 6000 & 120 \leq X \end{cases}$$

Solución:

a) Sea $X :=$ Resistencia a la ruptura. entonces $X \sim N(\mu, \sigma)$ Del enunciado tenemos $P[X < 114] = 0.8413$ y $P[X \geq 118] = 0.0228$. Esto es equivalente a:

$$P\left[Z \leq \frac{114-\mu}{\sigma}\right] = 0.8413, \qquad P\left[Z > \frac{118-\mu}{\sigma}\right] = 0.0228$$

Revizando la tabla podemos ver que:

$$\frac{114-\mu}{\sigma} = 1, \qquad \frac{118-\mu}{\sigma} = 2$$

Así, $\mu = 110$ y $\sigma^2 = 16$

b) Notemos que la función de cuantía de Y es:

$$f_Y(y) = \begin{cases} P[X \leq 105] & \text{si } y = 1000 \\ P[105 \leq X \leq 120] & \text{si } y = 3000 \\ P[120 \leq X] & \text{si } y = 6000 \end{cases} = \begin{cases} 0.1056 & \text{si } y = 1000 \\ 0.8881 & \text{si } y = 3000 \\ 0.0062 & \text{si } y = 6000 \end{cases}$$

De acá podmeos ver que $E[Y] = 2807$, por lo que la utilidad es $E[0.15 \cdot Y] = 0.15 \cdot 2807 = 421.1$

Distribuciones Adicionales - t-student

Una variable aleatoria X sigue una distribución t de Student con parámetro $\nu > 0$ (denominado "grados de libertad") si su función de densidad está dada por

$$f(x) = \frac{\Gamma(\nu/2 + 1/2)}{\Gamma(\nu/2)\sqrt{\nu\pi}} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \qquad x \in \mathbb{R}.$$

Usaremos la notación $X \sim \mathrm{t}(\nu)$.

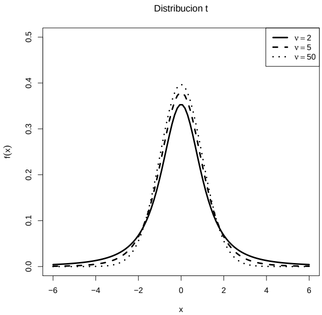


Figura 3: Densidad t-student con diferentes parámetros

Distribuciones Adicionales - Chi-cuadrado

Una variable aleatoria X sigue una distribución chi-cuadrado con parámetro $\nu \in \mathbb{N}$ (denominado "grados de libertad") si su función de densidad está dada por

$$f(x) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}}\, x^{\nu/2-1}\, e^{-x/2}, \qquad x > 0.$$

Usaremos la notación $X \sim \chi^2(\nu)$.

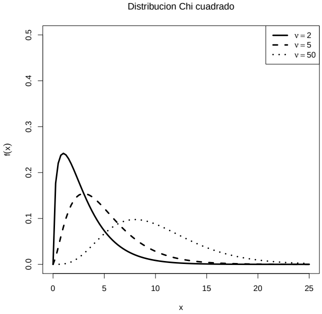


Figura 4: Densidad t-student con diferentes parámetros