



Unidad 04: Repaso de probabilidad y estadística.

Aprendizaje Automático

Docentes:
Diego P. Durante
ddurante@frba.utn.edu.ar

Ramiro Verrastro gramiro verrastro gramiro verrastro gramino del control del c

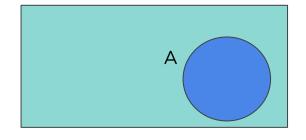
Tema 1 - Repaso de Probabilidad

Definiciones

- Conjunto
- Experimento
- Espacio muestral
- Suceso/Evento

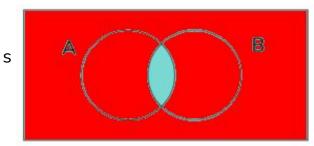


 $s=\{1,2,3,4,5,6\}$



S

Operaciones entre sucesos



- Unión
- Intersección
- Complemento
- Leyes de Demorgan

$$A \cup B$$

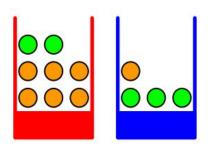
$$A \cap B$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

Probabilidad

- Definición
- Propiedades
- Probabilidad condicional



$$P(A) = cte. \in \Re$$

$$0 \le P(A) \le 1$$

$$P(s) = 1$$

$$P(B \cup A) = P(B) + P(A) - P(B \cap A)$$

$$P(B \cap A) = P(B) * P(A) \Leftrightarrow A y B son independientes$$

$$P(B \cap A) = P(B) * P(A|B)$$

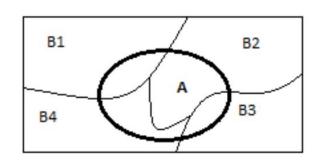
P(A): Probabilidad de elegir una pelota naranja

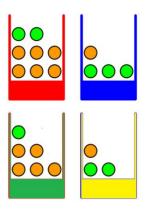
P(B): Probabilidad de elegir el cajón rojo

P(A|B): probabilidad de elegir una naranja, dado que elegí el cajón rojo.

 $P(B_r \cap A) = P(B_r) * P(A|B_r)$

Teorema de Bayes





$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^K P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^K P(B_i)P(A|B_i)}$$

para r = 1, 2, ..., k

Permutaciones y Combinatoria

$$n! = \prod_{k=1}^n k$$
 .

$$n! = n(n-1)(n-2)\cdots 1$$

$$C_n^p = inom{n}{p} = rac{n!}{p!(n-p)!}$$

n: cantidad de objetos

p: cantidad de elementos de los grupos.

Variables aleatorias

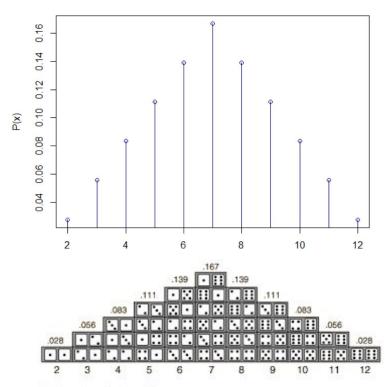
- Definición
- Tipos
 - o V.D.
 - o V.C.





Función de probabilidad variable discreta

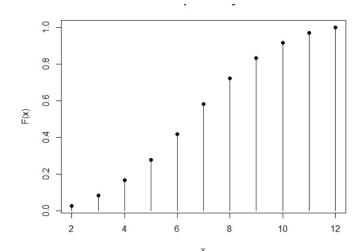
$$f(x) = P(X = x)$$
$$(x, f(x))$$
$$f(x) \ge 0$$
$$\sum f(x) = 1$$



Total number of states: 36

Función de distribución o probabilidad acumulada (V.D)

$$F(x) = P(X \le x) = \sum_{t \le x} f(t)$$
 para $-\infty < x < \infty$

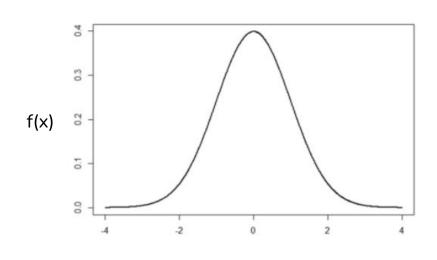


Función de densidad de probabilidad Variable continua

$$f(x) \ge 0$$

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

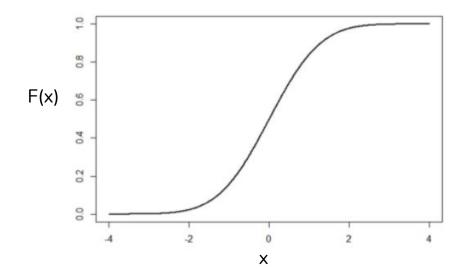
$$P(A < x < B) = \int_{A}^{B} f(x) dx$$



X

Función de distribución o probabilidad acumulada (V.C)

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$
 para $-\infty < x < \infty$



Preguntas?



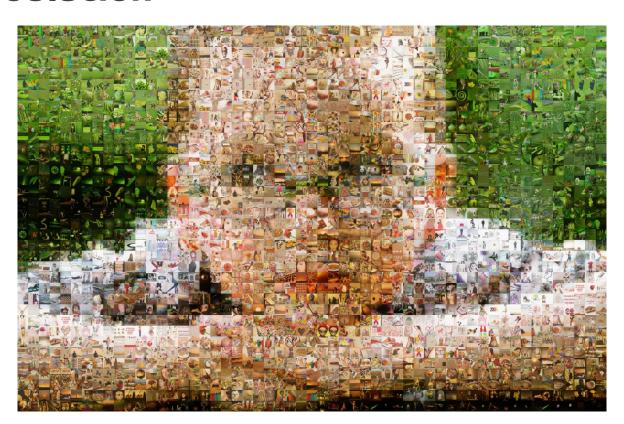
- Tema 1 Repaso de Probabilidad
- Tema 2 Repaso de Estadística

Estadística

Abarca métodos para:

- Recoger, organizar, resumir y analizar datos
- Sacar conclusiones válidas
- Tomar decisiones razonables basadas en tal análisis.

Población



Muestra



Muestreo aleatorio

E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24
E25	E26	E27	E28	E29	E30	E31	E32	E33	E34	E35	E36	E37	E38	E39	E40	E41	E42	E43	E44	E45	E46	E47	E48
E49	E50	E51	E52	E53	E54	E55	E56	E57	E58	E59	E60	E61	E62	E63	E64	E65	E66	E67	E68	E69	E70	E71	E72
E73	E74	E75	E76	E77	E78	E79	E80	E81	E82	E83	E84	E85	E86	E87	E88	E89	E90	E91	E92	E93	E94	E95	E96

E1	E2	E3	E4	E 5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24
E25	E26	E27	E28	E29	E30	E31	E32	E33	E34	E35	E36	E37	E38	E39	E40	E41	E42	E43	E44	E45	E46	E47	E48
E49	E50	E51	E52	E53	E54	E55	E56	E57	E58	E59	E60	E61	E62	E63	E64	E65	E66	E67	E68	E69	E70	E71	E72
E73	E74	E75	E76	E77	E78	E79	E80	E81	E82	E83	E84	E85	E86	E87	E88	E89	E90	E91	E92	E93	E94	E95	E96

Muestras aleatorias

E75 E32 E84 E66 E21

Muestreo por lotes

E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24
E25	E26	E27	E28	E29	E30	E31	E32	E33	E34	E35	E36	E37	E38	E39	E40	E41	E42	E43	E44	E45	E46	E47	E48
E49	E50	E51	E52	E53	E54	E55	E56	E57	E58	E59	E60	E61	E62	E63	E64	E65	E66	E67	E68	E69	E70	E71	E72
E73	E74	E75	E76	E77	E78	E79	E80	E81	E82	E83	E84	E85	E86	E87	E88	E89	E90	E91	E92	E93	E94	E95	E96

E34	E37	E31	E28	E66	E70	E59	E49	E67	E56	E44	E45	E38	E32	E36	E30	E39	E41	E26	E42	E46	E40	E77	E75
E2	E11	E3	E21	E24	E23	E9	E18	E96	E95	E22	E16	E89	E74	E83	E73	E35	E25	E43	E13	E7	E4	E82	E10
E94	E88	E92	E14	E15	E17	E64	E68	E65	E50	E69	E62	E81	E90	E53	E51	E91	E80	E84	E78	E60	E54	E58	E33
E1	E19	E8	E12	E6	E20	E5	E48	E47	E72	E71	E57	E61	E55	E85	E79	E76	E93	E86	E87	E29	E27	E52	E63

Muestreo aleatorizado por lotes

E43	E25	E41	E35	E40	E26	E46	E31	E44	E38	E30	E37	E45	E39	E36	E48	E29	E42	E32	E28	E34	E47	E33	E27
E91	E73	E89	E83	E88	E74	E94	E79	E92	E86	E78	E85	E93	E87	E84	E96	E77	E90	E80	E76	E82	E95	E81	E75
E67	E49	E65	E59	E64	E50	E70	E55	E68	E62	E54	E61	E69	E63	E60	E72	E53	E66	E56	E52	E58	E71	E57	E51
E19	E1	E17	E11	E16	E2	E22	E7	E20	E14	E6	E13	E21	E15	E12	E24	E5	E18	E8	E4	E10	E23	E9	E3

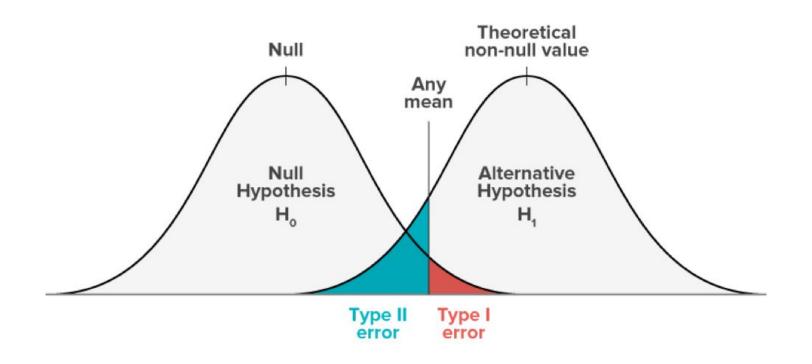
E34	E37	E31	E28	E66	E70	E59	E49	E67	E56	E44	E45	E38	E32	E36	E30	E39	E41	E26	E42	E46	E40	E77	E75
E2	E11	E 3	E21	E24	E23	E9	E18	E96	E95	E22	E16	E89	E74	E83	E73	E35	E25	E43	E13	E7	E4	E82	E10
E94	E88	E92	E14	E15	E17	E64	E68	E65	E50	E69	E62	E81	E90	E53	E51	E91	E80	E84	E78	E60	E54	E58	E33
E1	E19	E8	E12	E6	E20	E5	E48	E47	E72	E71	E57	E61	E55	E85	E79	E76	E93	E86	E87	E29	E27	E52	E63

Ramas de la estadística

Estadística descriptiva



Estadística inferencial



Estadísticos

Media y Esperanza

$$\mathrm{E}[X] = \sum_{i=1}^n x_i \, \mathrm{P}[X = x_i]$$

$$\mathrm{E}[X] = \int_{\scriptscriptstyle{\mathbb{D}}} x f_X(x) dx$$

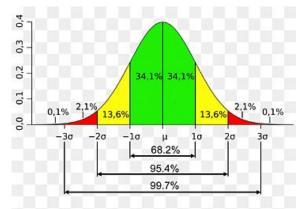
$$\mu = E(X) = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + \dots + x_n \cdot f(x_n) = \sum_{i=1}^n x_i \cdot f(x_i)$$

Varianza

$$\sigma^2 = V(X) = \sum (x - \mu)^2 \cdot f(x)$$

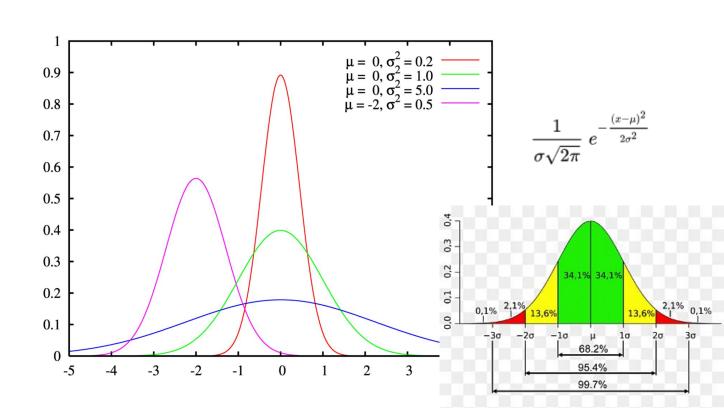
Desvío estándar

$$\sigma = \sqrt{\sigma^2}$$

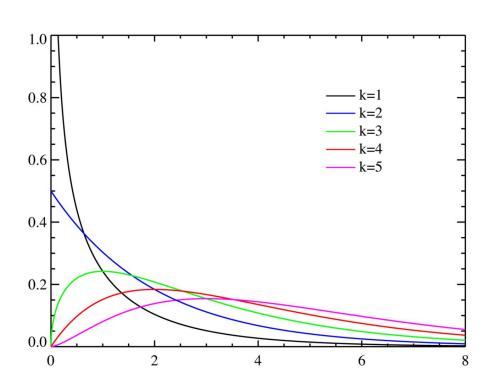


Distribución muestral

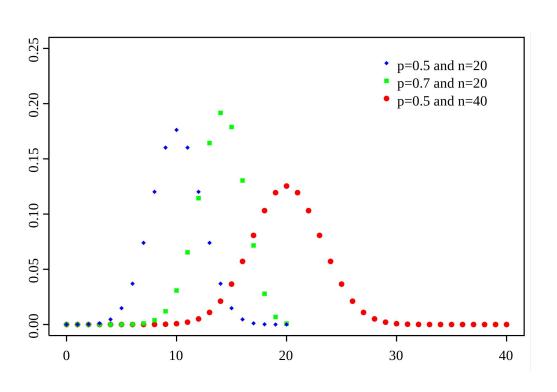
Distribución Normal



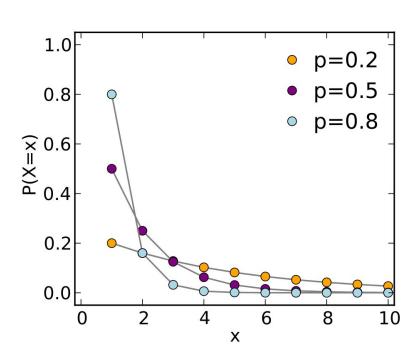
Distribución Chi-cuadrado



Distribución Binomial



Distribución Geométrica



Distribuciones de variable contínua

		Variables Aleate	orias Continuas	
$Distribuci\'on$	$f\left(x\right)$	$F\left(x\right)$	$E\left(x\right)$	Var(x)
$Uniforme\left(a,b ight)$	$\frac{1}{b-a}$	$\begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x \geq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$Exponencial(\lambda)$	$\lambda \ e^{-\lambda x}$	$1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$Normal\left(\mu,\sigma^2\right)$	$\frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma}$	$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$	μ	σ^2
$Gamma\ (k, \theta)$	$x^{k-1} \frac{e^{-x/\theta}}{\Gamma(k) \theta^k}$	$\int_0^x f(u;k,\theta)du$	$k\theta$	$k\theta^2$
$Beta(\alpha, \beta)$	$rac{x^{lpha-1}\left(1-x ight)^{eta-1}}{B\left(lpha,eta ight)}$	$I_{x}\left(lpha,eta ight)$	$\frac{\alpha}{\alpha + \beta}$	$\frac{\alpha\beta}{\left(\alpha+\beta\right)^2\left(\alpha+\beta+1\right)}$
$Weibull\left(\lambda,k ight)$	$\frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}$	$1 - e^{-(x/\lambda)^k}$	$\lambda \Gamma \left(1 + \frac{1}{k}\right)$	$\lambda^2 \Gamma \left(1 + \frac{2}{k}\right) - \mu^2$
$Cauchy\left(x_{0},\gamma ight)$	$\frac{1}{\pi\gamma\left(1+\left(\frac{x-x_0}{\gamma}\right)^2\right)}$	$\frac{1}{\pi} \tan^{-1} \left(\frac{x - x_0}{\gamma} \right) + \frac{1}{2}$	no existe	no existe
$Paretto\left(x_{m}, \alpha\right)$	$\frac{\alpha \ x_m^{\alpha}}{x^{\alpha+1}}, x > x_m$	$1-\left(\frac{x_m}{x}\right)^{\alpha}$	$\frac{\alpha \ x_m}{\alpha - 1}, \alpha > 1$	$\frac{x_m^2 \alpha}{(\alpha - 1)^2 (\alpha - 2)}, \alpha > 2$

Distribuciones de variable discreta

		Variables Aleator	ias Discretas	
Distribución	f(x)	$F\left(x\right)$	E(x)	Var(x)
$Bernoulli\left(p\right)$	p^xq^{1-x}	0 si x < 0 q si 0 < x < 1 I si x > I	p	pq
Binomial(n, p)	$\binom{n}{x} p^x q^{n-x}$	$\sum_{i=0}^{x} \binom{n}{i} p^i q^{n-i}$	np	npq
$Poisson\left(\lambda\right)$	$\frac{\lambda^x e^{-\lambda}}{x!}$	$e^{-\lambda} \sum_{i=0}^x \frac{\lambda^i}{i!}$	λ	λ
Geométrica I (p)	pq^{x-1}	$1-q^x$	$\frac{1}{p}$	$\frac{q}{p^2}$
Geométrica II (p)	pq^x	$1 - q^{x+1}$	$\frac{q}{p}$	$\frac{q}{p^2}$
$BinNeg\ I\ (r,p)$	$\binom{x-1}{r-1}p^rq^{x-r}$	$1-I_p(x+1,r)$	$\frac{r}{p}$	$\frac{rq}{p^2}$
$BinNeg\ II\ (r,p)$	$\binom{x+r-1}{r-1}p^rq^x$	$1-I_p(x+1,r)$	$\frac{rq}{p}$	$\frac{rq}{p^2}$
HiperGeo(N, m, n)	$\frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$	$\sum_{i=0}^{x} \frac{\binom{m}{i} \binom{N-m}{n-i}}{\binom{N}{n}}$	$\frac{nm}{N}$	$\frac{nm}{N}\left(1-\frac{m}{N}\right)\left(\frac{N-m}{N-1}\right)$

Preguntas?



