Instituto Tecnológico Autónomo de México

Repaso Probabilidad

Luis Gerardo Martínez Valdés
Mail: luis.martinez@neverlandconsultants.capital

Otoño 2021

Índice general

Yo sugeriría dejarlo como esto o bien ~~si~~ decir si $\{\omega \in \Omega \mid \mathcal{X}(\omega) = x\} \in \mathcal{F}$ $\forall x \in \mathbb{R}$ pero no hablar de Borelianos.

1. Variables Aleatorias

Definición 1 (Variable Aleatoria)**.** *Una función ~~$\mathcal{X} : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es variable aleatoria (v.a)~~ si ~~$\forall B \in \mathcal{B}(\mathbb{R})$ $\mathcal{X}^{-1} \in \mathcal{F}$ [2]~~*

*P~~ara fines de este curso diremos que~~ una **v.a.** es una función que va del espacio muestral Ω a los reales $\mathbb{R}$; definida por $\mathcal{X} : \Omega \longrightarrow \mathbb{R}$.*

Definición 2 (Funcion de Masa de Probabilidad (f.m.p))**.** *Si $\mathbb{P}(X = x) = \mathbb{P}\{\omega \in \Omega \ : \ X(\omega) = x\}$ es una f.m.p, entonces:*

llaves

Aclarar $\mathcal{X}$ discreta.

- $0 \leq \mathbb{P}(X = x) \leq 1$
- $\sum_{x \in S} \mathbb{P}(X = x) = 1$

donde S se conoce como soporte; $S = \{x \in \mathbb{R} \mid \mathbb{P}(\mathcal{X} = x) > 0\}$. en caso discreto.

decir y denota $f_{\mathcal{X}}(x)$

Definición 3 (Funcion de Densidad de Probabilidad (f.d.p))**.** *Si $f(x)$ es una f.d.p, entonces:*

- $0 \leq f(x) \leq 1$
- $\int_{-\infty}^{\infty} f(x)\, dx = 1$

También, se tiene que $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)\, dx$.

Definición 4 (Función de Distribución Acumulada)**.** *Sea $\mathcal{X}$ una v.a. la F.d.a* cumple:

mayúscula

$$F_{\mathcal{X}} : \mathbb{R} \longrightarrow [0, 1], \ F_{\mathcal{X}}(x) = \mathbb{P}(X \leq x)$$

- *No-Decreciente: Si $x_1 < x_2$ entonces $F(x_1) < F(x_2)$*
- *Normalizada: $F(-\infty) = 0$ y $F(\infty) = 1$*
- *Continua por la derecha: $\lim_{y \to x^+} F(y) = F(x)$*

falta con límites por la izq.

- *Si X es continua, entonces*

No hace falta parcial

$$\frac{dF(x)}{dx} = f(x); \quad F(x) = \int_{-\infty}^{x} f(t)\, dt$$

Definición 5 (Función de Densidad de Probabilidad **Condicional**)**.**

$$\mathbb{P}(a \leq Y \leq b \mid X = x) = \int_a^b f_{Y|X}(y|x)\, dy \ ; \ a \leq b$$

$$\cancel{\int_a^b} f_{Y|X}(y|x) \cancel{dy} = \frac{f(x, y)}{f_X(x)}$$

Donde $f(x, y)$ es la distribución conjunta de X, Y y $f_X(x)$ es la funsión de densidad marginal de X.

$$f_{\mathcal{X}}(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

Sean X, Y v.a.'s independientes, entonces se tiene que:

1. $F(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) = F(x)F(y)$
2. $f_{X,Y}(x,y) = f_X(x)f_Y(y)$
3. $p(x,y) = p_X(x)\, p_Y(y)$

1.1 Esperanza, Varianza, Covarianza y Correlación

[1] Definiciones y propiedades, sean X, Y, Z va.a's:

1.1.1 Esperanza

1. $\mathbb{E}(X) = \mu_X = \begin{cases} \displaystyle\sum_{x \in S} x f(x) & X \text{ es discreta} \\ \displaystyle\int_{-\infty}^{\infty} x f_X(x)\, dx & X \text{ es continua} \end{cases}$
2. $\mathbb{E}(cX) = c\,\mathbb{E}(X)\ ;\ c \in \mathbb{R}$
3. $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
4. $\mathbb{E}(g(X,Y)) = \int_{X,Y} g(x,y) f_{X,Y}(x,y)\, dx\, dy$ para cualquier función real valuada $g(\bullet)$
5. $\mathbb{E}(X) = \sum_{x=1}^{\infty} \mathbb{P}(X \geq x)$ con X discreta
6. Esperanza Condicional: $\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y f(y|x)\, dy$
7. Ley de Esperanza Total(Teorema de Torre): $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y))$
8. $\mathbb{E}(g(X,Y)|X = x) = \int_{-\infty}^{\infty} g(x,y) f_{Y|X}(y|x)\, dx$
9. $\mathbb{E}(g(Y,Z)|X = x) = \int_{-\infty}^{\infty} g(y,z) f_{Y,Z|X}(y,z|x)\, dydz$ creo se volteó el condicional.
10. $\mathbb{E}(X + Y|X) = \mathbb{E}(Y|X) + \mathbb{E}(Z|X)$?
11. $\mathbb{E}(g(X)Y|X) = g(X)\mathbb{E}(Y|X)$

Se recomienda al lector probar algunas de las propiedades anteriores, en especial: 6,7,10 y 11.

1.1.2 Varianza

1. $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$
2. $\mathbb{V}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathbb{V}(X_i) + \sum_{i \neq j} COV(X_i, X_j) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2\sum_{i<j} COV(X_i, X_j)$
3. Varianza Condicional: $\mathbb{V}(Y|X) = \mathbb{E}(Y^2|X) - \mathbb{E}^2(Y|X)$
4. Varianza Total: $\mathbb{E}(\mathbb{V}(Y|X)) + \mathbb{V}(\mathbb{E}(Y|X))$

Se recomienda al lector demostrar las propiedades 2 y 4.

1.1.3 Covarianza

1. $COV(X,Y) = \mathbb{E}\big[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\big]$
2. $COV(X,y) = E(XY) - E(X)E(Y)$
3. $COV(X,a) = 0$ donde $a \in \mathbb{R}$
4. $COV(X,X) = \mathbb{V}(X)$
5. $COV(aX,bY) = ab\; COV(X,Y)$ donde $a,b \in \mathbb{R}$
6. $COV(X+a, Y+b) = COV(X,Y)$
7. $COV(\sum X_i, \sum Y_i) = \sum\sum COV(X_i, Y_i)$

1.1.4 Correlación

1. $Corr(X,Y) = \rho(X,Y) = \frac{COV(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}$
2. $\rho(c,Y) = 0 \;\; \forall \;\; c \in \mathbb{R}$
3. $\rho(X,Y) \in \big[-1,1\big]$
4. Sea $Y = aX + b$ entonces $\rho(X,Y) = \begin{cases} \rho(X,Y) = 1 & si\ a > 0 \\ \rho(X,Y) = -1 & si\ a < 0 \end{cases}$

Se recomienda al lector demostrar las propiedades 3 y 4.

1.1.5 Desigualdades Importantes

1. Desigualdad de Markov Si X es una v.a. y $g(\bullet)$ es una funcion tal que $g(x) \geq 0 \;\forall\; x$ entonces $\forall k$ se tiene que

$$\mathbb{P}(g(X) \geq k) \leq \frac{\mathbb{E}(g(x))}{k} \tag{1.1}$$

2. Desigualdad de Tchebysheff

$$\mathbb{P}\left(\left|\frac{x-\mu}{\sigma}\right|^2 \geq t^2\right) \leq \frac{1}{t^2} \tag{1.2}$$

Otras versiones:

$$\mathbb{P}(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

$$\mathbb{P}(|x-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

3. Desigualdad de Jensen: Sea $g : \mathbb{R} \longrightarrow \mathbb{R}$ una función, tal que

$$\begin{cases} \mathbb{E}(g(X)) \geq g(E(X)) & \text{g convexa} \\ \mathbb{E}(g(X)) \leq g(E(X)) & \text{g cóncava} \end{cases} \tag{1.3}$$

2. Distrbuciones Conocidas

Acá les pasé el formulario de Casella & Berger te lo mando

2.1 Continuas

	Notation	PDF	CDF	Expectation	Variance
Uniform	$U(a,b)$	$\frac{1}{b-a}\ x \in (a,b)$	$\frac{x-a}{b-a}\ x \in (a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	$\mathbb{N}(\mu,\sigma^2)$	$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-(x-\mu)^2/2\sigma^2\right)$	$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt$	μ	σ^2
Log-Normal	$\mathbb{LN}(\mu,\sigma^2)$	$\frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$	$\frac{1}{2}+\frac{1}{2}\text{erf}\left[\frac{\ln x-\mu}{\sqrt{2\sigma^2}}\right]$	$e^{\mu+\sigma^2/2}$	$(e^{\sigma^2}-1)e^{2\mu+\sigma^2}$
Student's t	Student(ν)	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}$	$I_x\left(\frac{\nu}{2},\frac{\nu}{2}\right)$	0	$\frac{\nu}{\nu-2};\ \nu>2$
Chi-square	$\chi^2(\nu)$	$\frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{\nu/2-1}e^{-x/2}$	$\frac{1}{\Gamma(\nu/2)}\gamma\left(\frac{\nu}{2},\frac{x}{2}\right)$	ν	2ν
Exponential	EXP(λ)	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
Gamma	Gam(α,β)	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	$\frac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Beta	Beta(α,β)	$\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$I_x(\alpha,\beta)\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{s^k}{k!}$
Weibull	Weibull(λ,k)	$\frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}\exp\left(\frac{x}{\lambda}\right)^k$	$1-e^{-(x/\lambda)^k}$	$\lambda\Gamma\left(1+\frac{1}{k}\right)$	$\lambda^2\Gamma\left(1+\frac{2}{k}\right)-\mu^2$
Pareto	Pareto(x_m,α)	$\frac{\alpha\, x_m^\alpha}{x^{\alpha+1}}\quad x \geq x_m$	$1-\left(\frac{x_m}{x}\right)^\alpha\ x \geq x_m$	$\frac{\alpha x_m}{\alpha-1}\ \alpha>1$	$\frac{x_m^2\alpha}{(\alpha-1)^2(\alpha-2)}\ \alpha>2$

2.2 Discretas

	Notation	PDF	CDF	Expectation	Variance	MGF
Bernoulli	$\mathrm{Ber}(p)$	$p^x\,(1-p)^{1-x}$	$(1-p)^{1-x}$	p	$p(1-p)$	$1-p+pe^s$
Binomial	$\mathrm{Bin}(n,p)$	$\binom{n}{x}p^x\,(1-p)^{n-x}$	$--$	np	$np(1-p)$	$(1-p+pe^t)^n$
Multinomial	$\mathrm{Multi}(n,p_1,\cdots,p_n)$	$\frac{n!}{x_1!\cdots x_k!}p_1^{x_1}\cdots p_k^{x_k} \quad \sum_{i=1}^{k} x_i = n$	$--$	$\begin{pmatrix} np_1 \\ \vdots \\ np_k \end{pmatrix}$	$\begin{pmatrix} np_1(1-p_1) & -np_1p_2 \\ -np_2p_1 & \ddots \end{pmatrix}$	$\left(\sum_{i=0}^{k} p_i e^{t_i}\right)^n$
Geometric[1]	$\mathrm{Geo}(p)$	$p(1-p)^{x-1}$	$1-(1-p)^x$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hypergeometric	$\mathrm{HG}(N,m,n)$	$\frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$	$--$	$\frac{nm}{N}$	$\frac{nm(N-n)(N-m)}{N^2(N-1)}$	
Negative Binomial	NB	$\binom{x+r-1}{r-1}p^r(1-p)^x$	$--$	$r\,\frac{1-p}{p}$	$r\,\frac{1-p}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Poisson	$\mathrm{Po}(\theta)$	$\frac{e^{-\theta}\theta^x}{x!}$	$--$	θ	θ	$e^{\theta(e^t-1)}$

Esto te creo hay varias que no salía

3. Relación entre distribuciones

3.1 Binomial

- $Bin(1, p) \equiv Ber(p)$
- Si $X_i \overset{v.a.i.i.d}{\sim} Bin(1, p) \Rightarrow \ \sum_{i=1}^{n} X_i \sim Bin(n, p)$
- Si $X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p)$ independientes, entonces $X_1 + X_2 \sim Bin(n_1 + n_2, p)$
- $Bin(n, p) \equiv Poisson(np)$ para n grande y p pequeña en límite.
- $Bin(n, p) \equiv N(np, np(1-p))$ para n grande y una p ni tan pequña ni tan grande ?

por aprox

3.2 Binomial Negativa

- $BN(1, p) \equiv Geo(p)$
- $X_i \overset{v.a.i.i.d}{\sim} Geo(p) \Rightarrow \ \sum_{i=1}^{n} X_i \sim BN(n, p)$
- Si $X_i \overset{v.a.i.i.d}{\sim} Geo(p) \Rightarrow \ X_{(1)} = min(X_1, ..., X_n) \sim \ Geo(1 - (1-p)^n)$

3.3 Poisson

- Si $X_i \overset{v.a.i.i.d}{\sim} Poisson(\lambda_i) \Rightarrow \ \sum_{i=1}^{n} X_i \sim Possion(\sum_{i=1}^{n} \lambda_i)$
- Si $X_i \overset{v.a.i.i.d}{\sim} Poisson(\lambda_i) \Rightarrow \ X_i | \sum_{j=1}^{n} X_j \sim Bin(\sum_{j=1}^{n} X_j, \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j})$

3.4 Exponencial

- Si $X_i \overset{v.a.i.i.d}{\sim} Exp(\lambda_i) \Rightarrow \ \sum_{i=1}^{n} X_i \sim Gamma(n, \ \sum_{i=1}^{n} \lambda_i)$
- Si $X_i \overset{v.a.i.i.d}{\sim} Exp(\lambda)$ entonces:
  1. $X_{(1)} \ \sim \ Exp(n\lambda)$
  2. $X^{\frac{1}{\alpha}} \sim Weibull(\alpha, \lambda)$

3.5 Uniforme

- Si $X_i \overset{v.a.i.i.d}{\sim} U(0, 1)$ entonces:
  1. $-\theta ln(X_i) \ \sim \ Exp(\frac{1}{\theta})$ donde θ es usualmente un parámetro que queremos estimar.
  2. $-2ln(X_i) \ \sim \ \chi^2_{(2)}$

3.6 Gamma

- Si $X_i \overset{v.a.i.i.d}{\sim} Gamma(r, \lambda_i) \;\Rightarrow\; \sum_{i=1}^{n} X_i \sim Gamma(r,\; \sum_{i=1}^{n} \lambda_i)$
- $Gamma(1, \lambda) \equiv Exp(\lambda)$
- Si $X \sim Gamma(r, \lambda)$ entonces para $k > 0,\; kX \;\sim Gamma(r, \frac{\lambda}{k})$
- $Gamma(\frac{\nu}{2}) \equiv \chi^2_{(\nu)}$
- $X_1 \sim Gamma(r_1, \lambda),\; X_2 \sim Bin(r_2, \lambda)$ independientes, entonces $\frac{X_1}{X_1+X_2} \sim Beta(r_1, r_2)$

3.7 Chi-Cuadrada

- $X_i \overset{v.a.i.i.d}{\sim} \chi^2_{(\nu_i)} \;\Rightarrow\; \sum_{i=1}^{n} X_i \sim \chi^2_{(\sum_{i=1}^{n} \nu_i)}$
- $X_1 \sim \chi^2_{(\nu_1)},\; X_2 \sim \chi^2_{(\nu_2)}$ independientes, entonces:
  1. $\frac{\frac{X_1}{\nu_1}}{\frac{X_2}{\nu_2}} \sim F(\nu_1, \nu_2)$
  2. $\frac{X_1}{X_1+X_2} \sim Beta(\frac{\nu_1}{2}, \frac{\nu_2}{2})$

3.8 Normal

- Si $Z_i \overset{v.a.i.i.d}{\sim} N(0,1) \;\Rightarrow\; \sum_{i=1}^{n} Z_i^2 \sim \chi^2_{(n)}$
- Si Z_1 es independiente de Z_2 entonces $\frac{Z_1}{Z_2} \sim Cauchy(0,1) \equiv t(1)$

3.9 t-Student

- Si $Z \sim N(0,1)$ es independiente de $V \sim \chi^2_{\nu}$ entonces $\frac{Z}{\sqrt{\frac{V}{\nu}}} \;\sim t(\nu)$
- Si $X \sim t(\nu) \Rightarrow\; X^2 \sim F(1, \nu)$

3.10 Pareto

- Si $X_i \overset{v.a.i.i.d}{\sim} Pareto(\alpha, \beta)$ con $i = 1, ..., n$ entonces:
  1. $ln(\frac{X_i}{\beta}) \sim Exp(\alpha)$
  2. $X_{(1)} \sim Pareto(n\alpha, \beta)$
  3. $2\alpha \sum_1^n (\frac{X_i}{\beta}) \sim \chi^2_{(2n)}$

3.11 Beta

- Si $X \sim Beta(\alpha, 1)$ entonces:
  1. $-ln(X) \sim Exp(\alpha)$
  2. $-2\alpha ln(X) \sim \chi^2_{(2)}$
- Si $X \sim Beta(1, \beta)$ entonces:
  1. $-ln(1-X) \sim Exp(\beta)$
  2. $-2\beta ln(1-X) \sim \chi^2_{(2)}$

4. Ejercicios

1. Una tarea llega a un sistema y es atendida por el procesador C_i con una probabilidad $p_i,\ i = 1, .., n$. El tiempo aleatorio que toma el procesador en terminar la tarea es distribuido exponencialmente con parámetro λ_i

   (a) Muestre que la función de densidad de $T \equiv$ el tiempo que toma el sistema para el procesamiento de la tarea está dada por:
   $$f_T(t) = \sum_{i=1}^{n} p_i \lambda_i e^{-\lambda_i t}, \quad 0 \leq t$$

   (b) Muestre entonces que:
   $$\mathbb{E}(T) = \sum_{i=1}^{n} \frac{p_i}{\lambda_i}$$

2. Demuestra el Teorema de Cambio de variable univariado para el caso estrictamente creciente i.e. Sea X una v.a. continua y ϕ una función estrictamente creciente. Entonces:
   $$f_{\phi(X)}(t) = f_X(\phi^{-1}(t)) \left| \frac{d}{dt} \phi^{-1}(t) \right|$$

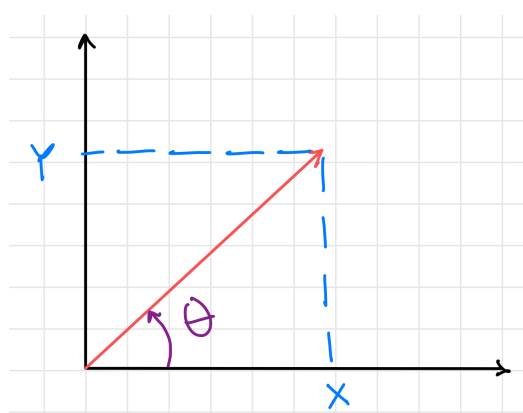
3. Para cierto tipo de suelo, el número de lombrices por pie cúbico tiene una media de 100. Suponiendo una distribución de Poisson de las lombrices, proponga un intervalo que incluirá al menos $\frac{5}{9}$ de los valores muestrales de las cantidades de lobrices ibtenidas de un número grande de muestras de 1 pie cúbico.

4. Si $X, Y \overset{v.a.i.i.d}{\sim} Exponenciales$ con parámetros λ_1, λ_2 respectivamente. Encuentra la f.d.p $Z = \frac{X}{Y}$ y calcule $\mathbb{P}(X < Y)$

5. Sea $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ valores ordenados de n v.a.'s $Uniformes(0, 1)$.
   Prueba para $1 < k < n + 1$ que $\mathbb{P}(X_{(k)} - X_{(k-1)} > t) = (1 - t)^n$ donde $X_{(0)} = 0,\ X_{(n+1)} = t$

6. Si $U_1, U_2 \sim Unif(0, 1)$ son independientes, encuentre la f.d.p de $|U_1 - U_2|$

7. Si $\theta \sim Unif(\frac{-\pi}{2}, \frac{\pi}{2})$. Encuentre la distribución de Y.



8. Demuestra la Desigualdad de Markov.

Ambas 8 y 9 veremos en clase pero está bien

9. Demuestra la Desigualdad de Tchebysheff

10. Sean $X_1, ..., X_n \overset{v.a.i.i.d.}{\sim} f(\bullet)$ tal que $\mathbb{E}(X_i) = \mu,\ \mathbb{V}(X_i) = \sigma^2$. Demuestra:

    (a) $\mathbb{E}(\bar{X}) = \mu$

    (b) $\mathbb{V}(\bar{X}) = \frac{\sigma^2}{n}$

    (c) $\mathbb{E}(S^2) = \sigma^2$

Bibliografía

[1] George Casella and Roger L Berger. *Statistical inference*. Cengage Learning, 2021.

[2] Dennis Wackerly, William Mendenhall, and Richard L Scheaffer. *Mathematical statistics with applications*. Cengage Learning, 2014.