Probability Notes

Mark DiBattista

March 11, 2019

Abstract

Probability.

1 Suggested Resource Materials

Useful source texts:

- Probability/Statistics, intermediate (probability sections are better than statistics): *Statistical Inference*, Casella & Berger
- Probability, advance: *Probability and Measure*, Billingsley

Throughout the text the acronym, *LAN*, refers to the companion writeup, *Linear Algebra Notes*, in which information is referenced by chapter and/or numbered equation.

2 Probability Preliminaries

$$\text{Probability: } \mathbb{P}E \tag{1}$$
$$\text{Mean: } \mathbb{E}X \tag{2}$$
$$\text{Variance: } \mathbb{V}X \tag{3}$$
$$\text{Covariance: } \mathbb{C}(X,Y) \tag{4}$$
$$\text{Information: } \mathbb{I}X \tag{5}$$
$$\text{Entropy: } \mathbb{H}X \tag{6}$$

3 Norms and Inequalities

3.1 Cauchy-Schwartz Inequality

$$|\langle \mathbf{x}_1, \mathbf{x}_2 \rangle|^2 \leq \langle \mathbf{x}_1, \mathbf{x}_1 \rangle \cdot \langle \mathbf{x}_2, \mathbf{x}_2 \rangle \tag{7}$$

$$\mathbb{C}(X_1, X_2) \leq \mathbb{V}X_1 \cdot \mathbb{V}X_2 \tag{8}$$

3.2 Chebyshev's Inequality

$$\mathbb{P}\{g(X) \geq r\} \leq \frac{\mathbb{E}g(X)}{r} = \int_D g(x)f(x)\,dx \geq \int_{\{g(x)\geq r\}} g(x)f(x)\,dx$$
$$\geq r\int_{\{g(x)\geq r\}} f(x)\,dx = r\mathbb{P}\{g(X)\geq r\}. \quad (9)$$

3.3 Jensen's Inequality

$$\phi(\mathbb{E}X) \leq \mathbb{E}\phi(X) \quad (10)$$

convex function

$$\phi(tx_1 + (1-t)x_2) \leq t\phi(x_1) + (1-t)\phi(x_2) \quad (11)$$

$$ax + b \leq \phi(x) \quad (12)$$

$$\mathbb{E}\phi(X) = \int_I \phi(x)p(x)\,dx \geq \int_I (ax+b)p(x)\,dx =$$
$$a\int_I xp(x)\,dx + b\int_I p(x)\,dx = ax_0 + b = \phi(x_0) = \phi(\mathbb{E}X). \quad (13)$$

4 Operators

4.1 Moment-Generating Functions

$$M_X(t) \equiv \mathbb{E}e^{tX} \quad (14)$$

$$M_X(t) \equiv \sum_n = 1^\infty \frac{\mathbb{E}X^n}{n!}t^n \quad (15)$$

$$\mathbb{E}X^n = \frac{d^n}{dt^n}M_X(t)|_{t=0} = M_X^{(n)}(0) \quad (16)$$

X and Y independent

$$M_{X+Y}(t) = \mathbb{E}e^{t(X+Y)} = \mathbb{E}e^{tX}e^{tY} = \mathbb{E}e^{tX}\mathbb{E}e^{tY} = M_X(t)M_Y(t) \quad (17)$$

$$M_{cX}(t) = \mathbb{E}e^{ctX} = M_X(ct) \quad (18)$$

4.2 Cumulants

$$K_X(t) = \ln M_X(t) \tag{19}$$

Since $\ln(1+x) = t - \frac{t^2}{2} + \cdots$

$$K_X(t) = \left(t\mathbb{E}X + \frac{t^2}{2}\mathbb{E}X^2 + \cdots\right) + \frac{1}{2}\left(t\mathbb{E}X + \cdots\right)^2 + \cdots \tag{20}$$

$$= t\mathbb{E}X + \frac{t^2}{2}\left((\mathbb{E}X)^2 - \mathbb{E}X^2\right) + \cdots \tag{21}$$

For X and Y independent

$$K_{X+Y}(t) = K_X(t) + K_Y(t) \tag{22}$$

$$K_{cX}(t) = K_X(ct) \tag{23}$$

4.3 Characteristic Functions

Fourier transform of probability density function

$$\phi_X(t) \equiv \mathbb{E}e^{itX} \tag{24}$$

for X and Y independent

$$\phi_{X+Y}(t) = \mathbb{E}e^{it(X+Y)} = \mathbb{E}e^{itX}e^{itY} = \mathbb{E}e^{itX}\mathbb{E}e^{itY} = \phi_X(t)\phi_Y(t) \tag{25}$$

$$\phi_{cX}(t) = \mathbb{E}e^{ictX} = \phi_X(ct) \tag{26}$$

let $Z = X + Y$ be the sum of two independent random variables

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) \Rightarrow p_{X+Y}(z) = \int_{-\infty}^{\infty} p_X(x)p_Y(z-x)\,dx \tag{27}$$

4.4 Extensions to Random Vectors

$$\left.\begin{matrix} X \to \mathbf{X} \\ t \to \mathbf{t} \end{matrix}\right\} \Rightarrow \begin{cases} M_{\mathbf{X}}(\mathbf{t}) \equiv \mathbb{E}e^{\mathbf{t}^\top \mathbf{X}} \\ K_{\mathbf{X}}(\mathbf{t}) \equiv \ln M_{\mathbf{X}}(\mathbf{t}) \\ \phi_{\mathbf{X}}(\mathbf{t}) \equiv \mathbb{E}e^{i\mathbf{t}^\top \mathbf{X}} \end{cases} \tag{28}$$

4.5 Transformations

$$Y = g(X) \tag{29}$$

$$\text{increasing function, } g: F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq Y = \mathbb{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y)) \tag{30}$$

$$\text{decreasing function, } g: F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq Y = \mathbb{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)) \tag{31}$$

$$f_Y(y) = \frac{d}{dy}F_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| \tag{32}$$

4.6 Scale-location Adjustment

$$X \sim f(x) \Rightarrow \alpha + \beta X \sim \frac{1}{\beta} f(\alpha + \beta x) \tag{33}$$

5 Common Functions

- Error function and complementary error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt \tag{34}$$

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) \tag{35}$$

- Gamma function

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \tag{36}$$

$$\begin{aligned}
\Gamma(x) &= \int_0^\infty t^{x-1} e^{-t}\, dt \\
&= -t^{x-1} e^{-t}\big|_0^\infty - \int_0^\infty (x-1) t^{x-2} (-e^{-t})\, dt \\
&= (x-1) \int_0^\infty t^{x-2} e^{-t}\, dt \\
&= (x-1)\Gamma(x-1)
\end{aligned} \tag{37}$$

$$\Gamma(n) = (n-1)! \tag{38}$$

- Beta function

$$\mathrm{B}(x,y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt \tag{39}$$

$$\begin{aligned}
\Gamma(x)\Gamma(y) &= \int_0^\infty s^{x-1} e^{-s}\, ds \int_0^\infty t^{y-1} e^{-t}\, dt \\
&= \int_0^\infty \int_0^\infty s^{x-1} t^{y-1} e^{-(s+t)}\, ds\, dt && \begin{cases} s = uv \\ t = u(1-v) \end{cases} \\
&= \int_{u=0}^\infty \int_{v=0}^1 (uv)^{x-1} (u(1-v))^{y-1} e^{-u} u\, du\, dv && |J| = u \\
&= \int_0^\infty e^{-u} u^{x+y-1}\, du \int_0^1 v^{x-1} (1-v)^{y-1}\, dv \\
&= \Gamma(x+y)\, \mathrm{B}(x,y)
\end{aligned} \tag{40}$$

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{41}$$

$$\begin{aligned}\mathrm{B}\left(\frac{1}{2},\frac{1}{2}\right) &= \int_0^1 t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\,dt \\ &= 2\int_0^{\frac{\pi}{2}} \frac{\cos\theta\sin\theta}{\cos\theta\sin\theta}\,d\theta \\ &= \pi \end{aligned} \tag{42}$$

$$\mathrm{B}\left(\frac{1}{2},\frac{1}{2}\right) = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)} \Rightarrow \Gamma\left(\frac{1}{2}\right) = \sqrt{\mathrm{B}\left(\frac{1}{2},\frac{1}{2}\right)} = \sqrt{\pi} \tag{43}$$

- Multivariate Beta Function

$$\mathrm{B}(\alpha_1,\ldots,\alpha_n) = \frac{\prod_{i=1}^n \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^n \alpha_i\right)} \tag{44}$$

$$\mathrm{B}(\alpha_1,\ldots,\alpha_n) = \frac{\Gamma(\alpha_j)\prod_{i\neq j}\Gamma(\alpha_i)}{\Gamma\left(\alpha_j + \sum_{i\neq j}\alpha_i\right)} = \frac{\prod_{i\neq j}\Gamma(\alpha_i)}{\Gamma(\sum_{i\neq j}\alpha_i)}\,\mathrm{B}\left(\alpha_j, \sum_{i\neq j}\alpha_i\right) \tag{45}$$

6 Common Distributions

6.1 Discrete Distributions

6.1.1 Bernoulli

$$\mathrm{Ber}(p) \equiv f(k|p) = p^k(1-p)^{1-k}, k \in \{0,1\} \tag{46}$$

6.1.2 Binomial

$$X_i \sim \mathrm{Ber}(p) \Rightarrow Y = \sum_{i=1}^n X_i \sim \mathrm{Bin}(n,p) \tag{47}$$

$$\mathrm{Bin}(n,p) \equiv f(k|n,p) = \binom{n}{k} p^k(1-p)^{n-k}, k \in \{0,\cdots,n\} \tag{48}$$

6.1.3 Negative Binomial

$$\mathrm{NB}(k|r,p) = \mathrm{Bin}(k|k+r-1,p)\,\mathrm{Ber}(0|p) \tag{49}$$

$$\mathrm{NB}(k|r,p) \equiv f(k|r,p) = \binom{k+r-1}{k} p^k(1-p)^{r-k}, k \in \mathbb{N} \tag{50}$$

6.1.4 Geometric

$$\mathrm{Geo}(k|p) = \mathrm{NB}(k|1, 1-p) = p(1-p)^{k-1}, k \in \mathbb{N} \tag{51}$$

6.1.5 Poisson

$$\mathrm{Poi}(\lambda) = f(k|\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}, k \in \mathbb{N} \tag{52}$$

$$\begin{aligned}
\mathrm{Bin}\left(n, \frac{\lambda}{n}\right) = f\left(k \,\middle|\, n, \frac{\lambda}{n}\right) &= \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{k}\right)^{n-k} \\
&= \frac{n \cdot n-1 \cdot \cdots \cdot n-k+1}{k!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^k \\
&= \left[\left(\frac{n}{n}\right)\cdot\left(\frac{n-1}{n}\right)\cdot\cdots\cdot\left(\frac{n-k+1}{n}\right)\right]\cdot\left[\left(1-\frac{\lambda}{n}\right)^{-k}\right]\cdot\left[\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\right]
\end{aligned} \tag{53}$$

$$\lim_{n\to\infty} \mathrm{Bin}\left(n, \frac{\lambda}{n}\right) = \frac{\lambda^k}{k!}e^{-\lambda} \equiv \mathrm{Poi}(\lambda) \tag{54}$$

6.1.6 Multinomial

The multinomial distribution is realized from the sum of n repeated *dependent* Bernoulli trials, each parametrized by potentially different probabilities of individual success, p_i, and linked by the requirement that one, and only one, may be successful on any given trial, $\sum_{i=1}^k p_i = 1$:

$$\mathrm{Mul}(n, p_1, \cdots, p_k) \equiv f(x_1, \cdots, x_k|n, p_1, \cdots, p_k) = \frac{n}{\prod_{i=1}^k x_i!}\prod_{i=1}^k p_i^{x_i} = \frac{\Gamma(1+\sum_{i=1}^k x_i)}{\prod_{i=1}^k \Gamma(1+x_i)}\prod_{i=1}^k p_i^{x_i} \tag{55}$$

$$X_1, \cdots, X_k \sim \mathrm{Mul}(n, p_1, \cdots, p_k) \Rightarrow \begin{cases} \mathbb{E}X_i = np_i \\ \mathbb{V}X_i = np_i(1-p_i) \\ \mathbb{C}(X_i, X_j) = -np_ip_j, i \neq j \end{cases} \tag{56}$$

6.1.7 Hypergeometric

$$\mathrm{Hyp}(n, N, K) \equiv f(k|n, N, K) = \frac{\binom{n}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \tag{57}$$

6.1.8 Multivariate Hypergeometric

$$\left.\begin{aligned} \mathbf{k} &= (k_1, \ldots, k_n)^\top \\ \mathbf{K} &= (K_1, \ldots, K_n)^\top \end{aligned}\right\} \Rightarrow \mathrm{MHG}(\mathbf{k}|\mathbf{K}) = \frac{\binom{K_1}{k_1}\cdots\binom{K_n}{k_n}}{\binom{\sum_{i=1}^n K_i}{\sum_{i=1}^n k_i}} \tag{58}$$

6.2 Continuous Distributions

$$f(x|\theta) = h(x)g(\theta)e^{\eta(\theta)T(x)} \tag{59}$$

6.2.1 Normal

6.2.1.1 Univariate Gaussian

$$\mathrm{N}(\mu,\sigma^2) \equiv f_{\mathrm{N}}(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{60}$$

$$\mathbb{P}_{\mathrm{N}(\mu,\sigma^2)}[-\infty,x] = F_{\mathrm{N}}(x|\mu,\sigma^2) = \int_{-\infty}^{x} f_{\mathrm{N}}(x|\mu,\sigma^2)\,dx = \frac{1}{2}\left(1+\operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right) \tag{61}$$

$$\mathbb{E}(X-\mu)^{2n-1} = 0 \tag{62}$$

$$g(\alpha) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\alpha\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sqrt{\alpha}} \tag{63}$$

$$\mathbb{E}(X-\mu)^{2n} = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^{2n}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = (-2\sigma^2)^n\,\left.\frac{d^n}{d\alpha^n}g(\alpha)\right|_{\alpha=1} = (2n-1)!!\,\sigma^{2n} \tag{64}$$

$$\begin{aligned}
M_{\mathrm{N}(\mu,\sigma^2)}(t) \equiv \mathbb{E}e^{tX} &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2-tx}{2\sigma^2}}\,dx && (65)\\
&= e^{t\mu+\frac{1}{2}t^2\sigma^2}\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+t\sigma^2))^2}{2\sigma^2}}\,dx && (66)\\
&= e^{t\mu+\frac{1}{2}t^2\sigma^2} && (67)
\end{aligned}$$

$$\phi_{\mathrm{N}(\mu,\sigma^2)}(t) \equiv \mathbb{E}e^{itX} = e^{it\mu-\frac{1}{2}t^2\sigma^2} \tag{68}$$

$$\left.\begin{aligned} X\sim\mathrm{N}(\mu,\sigma^2) &\Rightarrow M_X = e^{t\mu+\frac{1}{2}t^2\sigma^2}\\ Y\sim\mathrm{N}(\nu,\tau^2) &\Rightarrow M_Y = e^{t\nu+\frac{1}{2}t^2\tau^2}\end{aligned}\right\} \Rightarrow M_{X+Y} = e^{t(\mu+\nu)+\frac{1}{2}t^2(\sigma^2+\tau^2)} \Rightarrow X+Y\sim\mathrm{N}(\mu+\nu,\sigma^2+\tau^2) \tag{69}$$

6.2.1.2 Standard Normal

$$Z\sim\mathrm{N}(0,1) \tag{70}$$

6.2.1.3 Multivariate Gaussian

$$\mathbf{X} = \begin{pmatrix} X_1\\ \vdots\\ X_n\end{pmatrix} \Rightarrow \begin{cases} \boldsymbol{\mu} = \begin{pmatrix}\mathbb{E}X_1\\ \vdots\\ \mathbb{E}X_n\end{pmatrix}\\ \boldsymbol{\Sigma} = \begin{pmatrix}\mathbb{V}X_1 & \cdots & \mathbb{C}(X_1,X_n)\\ \vdots & \ddots & \vdots\\ \mathbb{C}(X_1,X_n) & \cdots & \mathbb{V}X_n\end{pmatrix}\end{cases} \tag{71}$$

$$\mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{\Sigma}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{72}$$

$$\mathbf{Z} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}) = (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\mathbf{z}^\top \mathbf{z}} \tag{73}$$

Transformation, $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{Z}$,

$$\left.\begin{aligned} \mathbb{E}\mathbf{X} &\equiv \mathbb{E}(\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{Z}) = \boldsymbol{\mu} \\ \mathbb{V}\mathbf{X} &\equiv \mathbb{E}\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)\left(\mathbf{X} - \mathbb{E}\mathbf{X}\right)^\top = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \end{aligned}\right\} \Rightarrow \mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top) \tag{74}$$

$$\boldsymbol{\Sigma} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \Rightarrow \boldsymbol{\Gamma} = \mathbf{Q}\mathbf{D}^{\frac{1}{2}}, \qquad \mathbf{D}^{\frac{1}{2}} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \tag{75}$$

Transformation, $\mathbf{Z} = \boldsymbol{\Gamma}^{-1}(\mathbf{X} - \boldsymbol{\mu})$

$$\left.\begin{aligned} &\mathbb{E}\mathbf{Z} \equiv \mathbb{E}\boldsymbol{\Gamma}(\mathbf{X} - \boldsymbol{\mu}) = \mathbf{0} \\ &\mathbb{V}\mathbf{Z} \equiv \mathbb{E}\left(\mathbf{Z} - \mathbb{E}\mathbf{Z}\right)\left(\mathbf{Z} - \mathbb{E}\mathbf{Z}\right)^\top = \mathbb{E}\mathbf{Z}\mathbf{Z}^\top = \boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-\top} = \mathbf{I} \end{aligned}\right\} \Rightarrow \mathbf{Z} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}) \tag{76}$$

$$\boldsymbol{\Sigma} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top \Rightarrow \boldsymbol{\Sigma}^{-1} = \mathbf{Q}\mathbf{D}^{-1}\mathbf{Q}^\top \tag{77}$$

$$\mathbf{D} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2) \Rightarrow \mathbf{D}^{-1} = \mathrm{diag}\left(\frac{1}{\sigma_1^2}, \ldots, \frac{1}{\sigma_n^2}\right) \tag{78}$$

$$\mathbf{y} = \mathbf{Q}^\top(\mathbf{x} - \boldsymbol{\mu}) \Rightarrow \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = \frac{1}{2}\mathbf{y}^\top \mathbf{D}^{-1}\mathbf{y} = \sum_{i=1}^n \frac{y_i^2}{2\sigma_i^2} \tag{79}$$

$$\mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^n \det \boldsymbol{\Sigma}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} = \prod_{i=1}^n \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{y_i^2}{2\sigma_i^2}} = \prod_{i=1}^n f_{\mathrm{N}}(y_i|0, \sigma_i^2) \tag{80}$$

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \Rightarrow \begin{cases} \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \\ \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \end{cases} \tag{81}$$

$$\mathbf{X}_1|(\mathbf{X}_2 = \mathbf{x}_2) \sim \mathrm{N}(\boldsymbol{x}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}, \tag{82}$$

6.2.2 Gamma

$$\Gamma(\alpha, \beta) \equiv f_\Gamma(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},\ x \geq 0 \tag{83}$$

$$\begin{aligned} \Gamma(\alpha) &= \int_0^\infty t^{\alpha-1} e^{-t}\, dt \\ &= \int_0^\infty (\beta x)^{\alpha-1} e^{-\beta x} \beta\, dx \qquad t = \beta x \\ &= \int_0^\infty \beta^\alpha x^{\alpha-1} e^{-\beta x}\, dx \end{aligned} \tag{84}$$

$$
\begin{aligned}
\mathbb{E}X^n &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^k x^{\alpha-1} e^{-\beta x}\, dx \\
&= \frac{\beta^\alpha}{\beta^{\alpha+k}} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \int_0^\infty \frac{\beta^{\alpha+k}}{\Gamma(\alpha+k)} x^{\alpha+k-1} e^{-\beta x}\, dx \\
&= \frac{1}{\beta^k} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \int_0^\infty f_\Gamma(x|\alpha+k,\beta)\, dx \\
&= \frac{1}{\beta^k} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \qquad (85)
\end{aligned}
$$

$$
\mathbb{E}X = \frac{\alpha}{\beta} \qquad (86)
$$

$$
\mathbb{V}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{(\alpha+1)\alpha}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2} \qquad (87)
$$

$$
\begin{aligned}
M_{\Gamma(\alpha,\beta)} \equiv \mathbb{E}e^{tX} &= \int_0^\infty e^{tx} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\beta-t)x}\, dx \\
&= \frac{\beta^\alpha}{(\beta-t)^\alpha} \int_0^\infty \frac{(\beta-t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\beta-t)x}\, dx \\
&= \left(\frac{\beta}{\beta-t}\right)^\alpha \int_0^\infty f_\Gamma(x|\alpha,\beta-t)\, dx \\
&= \left(\frac{\beta}{\beta-t}\right)^\alpha \qquad (88)
\end{aligned}
$$

$$
X_i \sim \Gamma(\alpha_i,\beta) \Rightarrow \sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n \alpha_i, \beta\right) \qquad (89)
$$

6.2.3 Inverse Gamma

6.2.4 Beta

$$
\mathrm{B}(\alpha,\beta) \equiv f_\mathrm{B}(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 \le x \le 1 \qquad (90)
$$

$$
\begin{aligned}
\mathbb{E}X^k &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^k x^{\alpha-1}(1-x)^{\beta-1}\, dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+\beta+k)} \int_0^1 \frac{\Gamma(\alpha+\beta+k)}{\Gamma(\alpha+k)\Gamma(\beta)} x^{\alpha+k-1}(1-x)^{\beta-1}\, dx \\
&= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+k)} \int_0^1 f_\mathrm{B}(x|\alpha+k,\beta)\, dx \\
&= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+k)} \qquad (91)
\end{aligned}
$$

$$\mathbb{E}X = \frac{\alpha}{\alpha+\beta} \tag{92}$$

$$\mathbb{V}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta}{(\alpha+\beta)^2} \tag{93}$$

6.2.5 Chi-square

$$X \sim \mathrm{N}(0,1) \Rightarrow X^2 \sim \chi^2 = \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \tag{94}$$

$$\mathbb{P}\{X^2 \le x\} = \mathbb{P}\{-\sqrt{x} \le X \le \sqrt{x}\} = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt \tag{95}$$

$$\begin{aligned}
\chi^2 \equiv f_{\chi^2}(x) &= \frac{d}{dx}\int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt \\
&= \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}}\right)\left(\frac{1}{2}x^{-\frac{1}{2}}\right) - \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}}\right)\left(-\frac{1}{2}x^{-\frac{1}{2}}\right) \\
&= \frac{1}{\sqrt{2\pi}} x^{\frac{1}{2}-1} e^{-\frac{x}{2}} \\
&= \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)
\end{aligned} \tag{96}$$

$$X_i \sim \mathrm{N}(0,1) \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right) \tag{97}$$

$$X \sim \chi_n^2 \Rightarrow \begin{cases} \mathbb{E}X = n \\ \mathbb{V}X = \frac{n}{2} \end{cases} \tag{98}$$

$$\mathbf{\Sigma} = \mathbf{QDQ}^\top = \mathbf{\Gamma\Gamma}^\top \Rightarrow \mathbf{Z} = \mathbf{\Gamma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \tag{99}$$

$$\mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \mathbf{\Sigma}) \Rightarrow (\mathbf{X} - \boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Z}^\top \mathbf{Z} \sim \chi_n^2 \tag{100}$$

6.2.6 F-Distribution

$$\begin{aligned}
&\left.\begin{matrix} X_1, \ldots, X_k \\ Y_1, \ldots, Y_m \end{matrix}\right\} \sim Z = \mathrm{N}(0,1) \Rightarrow \frac{\frac{1}{k}\sum_{i=1}^{k} X_i^2}{\frac{1}{m}\sum_{j=1}^{m} Y_j^2} \sim F(k,m), \\
&\qquad F(k,m) \equiv f_F(x|k,m) = \frac{\Gamma(\frac{k+m}{2})}{\Gamma(\frac{k}{2})\Gamma(\frac{m}{2})} \left(\frac{k}{m}\right)^{\frac{k}{2}} x^{\frac{k}{2}-1} \left(1 + \frac{k}{m}x\right)^{-\frac{k+m}{2}}
\end{aligned} \tag{101}$$

Define the numerator and denominator in terms of chi-square distributed variables

$$\left.\begin{array}{l} X=\sum_{i=1}^{k} X_i^2 \sim \chi_k^2 \\ Y=\sum_{j=1}^{m} Y_m^2 \sim \chi_m^2 \end{array}\right\} \Rightarrow \mathbb{P}\left\{\frac{\frac{X}{k}}{\frac{Y}{m}} \leq t\right\} = \mathbb{P}\{X \leq \frac{k}{m} tY\}$$

$$= \iint_{\{X \leq \frac{k}{m} tY\}} f_{\chi_k^2}(x) f_{\chi_m^2}(y)\, dx\, dy = \int_0^{\infty} \int_0^{\frac{k}{m} ty} f_{\chi_k^2}(x) f_{\chi_m^2}(y)\, dx\, dy \tag{102}$$

$$\begin{aligned} F(k,m) \equiv f_F(x|k,m) &= \frac{d}{dt} \int_0^{\infty} \int_0^{\frac{k}{m} ty} f_{\chi_k^2}(x) f_{\chi_m^2}(y)\, dx\, dy = \int_0^{\infty} \frac{d}{dt}\left(\int_0^{\frac{k}{m} ty} f_{\chi_k^2}(x)\, dx\right) f_{\chi_m^2}(y)\, dy \\ &= \frac{k}{m} \int_0^{\infty} f_{\chi_k^2}\left(\frac{k}{m} ty\right) f_{\chi_m^2}(y)\, y\, dy \\ &= \frac{k}{m} \int_0^{\infty} \left(\frac{\frac{1}{2}^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} \left(\frac{k}{m} ty\right)^{\frac{k}{2}-1} e^{-\frac{\frac{k}{m} ty}{2}}\right) \left(\frac{\frac{1}{2}^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)} y^{\frac{m}{2}-1} e^{-\frac{y}{2}}\right) y\, dy \\ &= \frac{k^{\frac{k}{2}}}{m} \int_0^{\infty} \frac{\left(\frac{1}{2}\right)^{\frac{k+m}{2}} t^{\frac{k}{2}-1}}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} y^{\frac{k+m}{2}-1} e^{-\frac{1}{2} y\left(\frac{k}{m} t+1\right)}\, dy \\ &= \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{k}{m}\right)^{\frac{k}{2}} t^{\frac{k}{2}-1} \left(\frac{1}{\frac{k}{m} t+1}\right)^{\frac{k+m}{2}} \int_0^{\infty} \frac{\left(\frac{t+1}{2}\right)^{\frac{k+m}{2}}}{\Gamma\left(\frac{k+m}{2}\right)} y^{\frac{k+m}{2}-1} e^{-y\frac{t+1}{2}}\, dy \\ &= \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{k}{m}\right)^{\frac{k}{2}} t^{\frac{k}{2}-1} \left(1+\frac{k}{m} t\right)^{-\frac{k+m}{2}} \int_0^{\infty} f_{\Gamma}\left(y \left| \frac{k+m}{2}, \frac{t+1}{2}\right.\right) dy \\ &= \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{k}{m}\right)^{\frac{k}{2}} t^{\frac{k}{2}-1} \left(1+\frac{k}{m} t\right)^{-\frac{k+m}{2}} \end{aligned} \tag{103}$$

6.2.7 T-Distribution

$$\left.\begin{array}{l} X_1 \\ Y_1, \ldots, Y_m \end{array}\right\} \sim Z = \mathrm{N}(0,1) \Rightarrow \frac{X_1}{\sqrt{\frac{1}{m}\sum_{j=1}^{m} Y_j^2}} \sim T(m),$$

$$T(m) \equiv f_T(x|m) = \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{m}{2})} \frac{1}{\sqrt{m}} \left(1+\frac{x^2}{m}\right)^{-\frac{m+1}{2}} \tag{104}$$

$$\left.\begin{array}{l} X = X_i^2 \sim \chi^2 \\ Y=\sum_{j=1}^{m} Y_m^2 \sim \chi_m^2 \end{array}\right\} \Rightarrow \mathbb{P}\left\{-t \leq \frac{\sqrt{X}}{\sqrt{\frac{1}{m} Y}} \leq t\right\} = \mathbb{P}\left\{\frac{X}{\frac{1}{m} Y} \leq t^2\right\}$$

$$= \mathbb{P}\left\{X^2 \leq \frac{t^2}{m} Y^2\right\} = \int_0^{\frac{t^2}{m}} f_F(x|1,m)\, dx \tag{105}$$

$$\begin{aligned} T(m) \equiv f_T(x|m) &= \frac{d}{dt} \int_0^{\frac{t^2}{m}} f_F(x|1,m)\, dx = \frac{2t}{m} f_F\left(\frac{t^2}{m} \middle| 1, m\right) \\ &= \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{m}{2}\right)} \frac{1}{\sqrt{m}} \left(1+\frac{t^2}{m}\right)^{-\frac{m+1}{2}} \end{aligned} \tag{106}$$

6.2.8 Cauchy

$$\left.\begin{matrix} U \\ V \end{matrix}\right\} \sim Z = \mathrm{N}(0,1) \Rightarrow \frac{U}{V} \sim \mathrm{Cau}(0,1) \equiv f_C(x) = \frac{1}{\pi}\frac{1}{x^2+1} \tag{107}$$

$$\begin{aligned}\mathbb{P}\left\{\frac{U}{V} \leq x\right\} = \mathbb{P}\left\{U \leq xV\right\} = \iint_{\{u \leq xv\}} f_{\mathrm{N}}(u|0,1) f_{\mathrm{N}}(v|0,1)\,du\,dv \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{xv} f_{\mathrm{N}}(u|0,1) f_{\mathrm{N}}(v|0,1)\,du\,dv\end{aligned} \tag{108}$$

$$\begin{aligned}\mathrm{Cau}(0,1) \equiv f_{\mathrm{C}}(x) &= \frac{d}{dx}\mathbb{P}\left\{\frac{U}{V} \leq x\right\} = \frac{d}{dx}\int_0^{\infty}\int_0^{xv} f_{\mathrm{N}}(u|0,1) f_{\mathrm{N}}(v|0,1)\,du\,dv \\ &= \int_{-\infty}^{\infty}\left(\frac{d}{dx}\int_{-\infty}^{xv} f_{\mathrm{N}}(u|0,1)\,du\right) f_{\mathrm{N}}(v|0,1)\,dv = \int_{-\infty}^{\infty} f_{\mathrm{N}}(xv|0,1) f_{\mathrm{N}}(v|0,1)\,dv \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2v^2}{2}} \frac{1}{2\pi} e^{-\frac{v^2}{2}} v\,dv = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{v^2(x^2+1)}{2}} v\,dv \\ &= \frac{1}{2\pi}\frac{1}{x^2+1}\int_{-\infty}^{\infty} e^{-t}\,dt = \frac{1}{\pi}\frac{1}{x^2+1}\int_0^{\infty} e^{-t}\,dt \\ &= \frac{1}{\pi}\frac{1}{x^2+1}\end{aligned} \tag{109}$$