- $\frac{\Gamma(\alpha)}{\lambda^\alpha} = \int_0^\infty x^{\alpha-1}e^{-\lambda x}\,dx$

Beta

- $\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$
- $\mathbb{E}\left[X^k\right] = \frac{\mathrm{B}(\alpha+k,\beta)}{\mathrm{B}(\alpha,\beta)} = \frac{\alpha+k-1}{\alpha+\beta+k-1}\mathbb{E}\left[X^{k-1}\right]$
- $\mathrm{Beta}\,(1,1) \sim \mathrm{Unif}\,(0,1)$

8 Probability and Moment Generating Functions

- $G_X(t) = \mathbb{E}\left[t^X\right] \qquad |t| < 1$
- $M_X(t) = G_X(e^t) = \mathbb{E}\left[e^{Xt}\right] = \mathbb{E}\left[\sum_{i=0}^\infty \frac{(Xt)^i}{i!}\right] = \sum_{i=0}^\infty \frac{\mathbb{E}\left[X^i\right]}{i!}\cdot t^i$
- $\mathbb{P}\left[X=0\right] = G_X(0)$
- $\mathbb{P}\left[X=1\right] = G'_X(0)$
- $\mathbb{P}\left[X=i\right] = \frac{G_X^{(i)}(0)}{i!}$
- $\mathbb{E}\left[X\right] = G'_X(1^-)$
- $\mathbb{E}\left[X^k\right] = M_X^{(k)}(0)$
- $\mathbb{E}\left[\frac{X!}{(X-k)!}\right] = G_X^{(k)}(1^-)$
- $\mathbb{V}\left[X\right] = G''_X(1^-) + G'_X(1^-) - \left(G'_X(1^-)\right)^2$
- $G_X(t) = G_Y(t) \implies X \stackrel{d}{=} Y$

9 Multivariate Distributions

9.1 Standard Bivariate Normal

Let $X, Y \sim \mathcal{N}(0,1) \wedge X \perp\!\!\!\perp Z$ where $Y = \rho X + \sqrt{1-\rho^2}Z$

Joint density

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right\}$$

Conditionals

$$(Y \mid X = x) \sim \mathcal{N}\left(\rho x, 1-\rho^2\right) \qquad \text{and} \qquad (X \mid Y = y) \sim \mathcal{N}\left(\rho y, 1-\rho^2\right)$$

Independence

$$X \perp\!\!\!\perp Y \iff \rho = 0$$

9.2 Bivariate Normal

Let $X \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$ and $Y \sim \mathcal{N}\left(\mu_y, \sigma_y^2\right)$.

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left\{-\frac{z}{2(1-\rho^2)}\right\}$$

$$z = \left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]$$

Conditional mean and variance

$$\mathbb{E}\left[X \mid Y\right] = \mathbb{E}\left[X\right] + \rho\frac{\sigma_X}{\sigma_Y}(Y - \mathbb{E}\left[Y\right])$$

$$\mathbb{V}\left[X \mid Y\right] = \sigma_X\sqrt{1-\rho^2}$$

9.3 Multivariate Normal

Covariance matrix Σ (Precision matrix Σ^{-1})

$$\Sigma = \begin{pmatrix} \mathbb{V}\left[X_1\right] & \cdots & \mathrm{Cov}\left[X_1, X_k\right] \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}\left[X_k, X_1\right] & \cdots & \mathbb{V}\left[X_k\right] \end{pmatrix}$$

If $X \sim \mathcal{N}(\mu, \Sigma)$,

$$f_X(x) = (2\pi)^{-n/2}\left|\Sigma\right|^{-1/2}\exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

Properties

- $Z \sim \mathcal{N}(0,1) \wedge X = \mu + \Sigma^{1/2}Z \implies X \sim \mathcal{N}(\mu, \Sigma)$
- $X \sim \mathcal{N}(\mu, \Sigma) \implies \Sigma^{-1/2}(X-\mu) \sim \mathcal{N}(0,1)$
- $X \sim \mathcal{N}(\mu, \Sigma) \implies AX \sim \mathcal{N}\left(A\mu, A\Sigma A^T\right)$
- $X \sim \mathcal{N}(\mu, \Sigma) \wedge \|a\| = k \implies a^TX \sim \mathcal{N}\left(a^T\mu, a^T\Sigma a\right)$

10 Convergence

Let $\{X_1, X_2, \ldots\}$ be a sequence of RV's and let X be another RV. Let F_n denote the CDF of X_n and let F denote the CDF of X.
Types of Convergence

1. In distribution (weakly, in law): $X_n \xrightarrow{\text{D}} X$

$$\lim_{n\to\infty} F_n(t) = F(t) \qquad \forall t \text{ where } F \text{ continuous}$$

2. In probability: $X_n \xrightarrow{\text{P}} X$

$$(\forall \varepsilon > 0) \; \lim_{n\to\infty} \mathbb{P}\left[|X_n - X| > \varepsilon\right] = 0$$

3. Almost surely (strongly): $X_n \xrightarrow{\text{as}} X$

$$\mathbb{P}\left[\lim_{n\to\infty} X_n = X\right] = \mathbb{P}\left[\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right] = 1$$

4. In quadratic mean (L_2): $X_n \xrightarrow{\text{qm}} X$

$$\lim_{n\to\infty} \mathbb{E}\left[(X_n - X)^2\right] = 0$$

Relationships

- $X_n \xrightarrow{\text{qm}} X \implies X_n \xrightarrow{\text{P}} X \implies X_n \xrightarrow{\text{D}} X$
- $X_n \xrightarrow{\text{as}} X \implies X_n \xrightarrow{\text{P}} X$
- $X_n \xrightarrow{\text{D}} X \wedge (\exists c \in \mathbb{R}) \, \mathbb{P}[X = c] = 1 \implies X_n \xrightarrow{\text{P}} X$
- $X_n \xrightarrow{\text{P}} X \wedge Y_n \xrightarrow{\text{P}} Y \implies X_n + Y_n \xrightarrow{\text{P}} X + Y$
- $X_n \xrightarrow{\text{qm}} X \wedge Y_n \xrightarrow{\text{qm}} Y \implies X_n + Y_n \xrightarrow{\text{qm}} X + Y$
- $X_n \xrightarrow{\text{P}} X \wedge Y_n \xrightarrow{\text{P}} Y \implies X_n Y_n \xrightarrow{\text{P}} XY$
- $X_n \xrightarrow{\text{P}} X \implies \varphi(X_n) \xrightarrow{\text{P}} \varphi(X)$
- $X_n \xrightarrow{\text{D}} X \implies \varphi(X_n) \xrightarrow{\text{D}} \varphi(X)$
- $X_n \xrightarrow{\text{qm}} b \iff \lim_{n\to\infty} \mathbb{E}[X_n] = b \wedge \lim_{n\to\infty} \mathbb{V}[X_n] = 0$
- $X_1, \ldots, X_n$ IID $\wedge\, \mathbb{E}[X] = \mu \wedge \mathbb{V}[X] < \infty \iff \bar{X}_n \xrightarrow{\text{qm}} \mu$

Slutzky's Theorem

- $X_n \xrightarrow{\text{D}} X$ and $Y_n \xrightarrow{\text{P}} c \implies X_n + Y_n \xrightarrow{\text{D}} X + c$
- $X_n \xrightarrow{\text{D}} X$ and $Y_n \xrightarrow{\text{P}} c \implies X_n Y_n \xrightarrow{\text{D}} cX$
- In general: $X_n \xrightarrow{\text{D}} X$ and $Y_n \xrightarrow{\text{D}} Y \not\Longrightarrow X_n + Y_n \xrightarrow{\text{D}} X + Y$

10.1 Law of Large Numbers (LLN)

Let $\{X_1, \ldots, X_n\}$ be a sequence of IID RV's, $\mathbb{E}[X_1] = \mu$.
Weak (WLLN)

$$\bar{X}_n \xrightarrow{\text{P}} \mu \qquad n \to \infty$$

Strong (SLLN)

$$\bar{X}_n \xrightarrow{\text{as}} \mu \qquad n \to \infty$$

10.2 Central Limit Theorem (CLT)

Let $\{X_1, \ldots, X_n\}$ be a sequence of IID RV's, $\mathbb{E}[X_1] = \mu$, and $\mathbb{V}[X_1] = \sigma^2$.

$$Z_n := \frac{\bar{X}_n - \mu}{\sqrt{\mathbb{V}\left[\bar{X}_n\right]}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\text{D}} Z \qquad \text{where } Z \sim \mathcal{N}(0,1)$$

$$\lim_{n\to\infty} \mathbb{P}[Z_n \le z] = \Phi(z) \qquad z \in \mathbb{R}$$

CLT notations

$$Z_n \approx \mathcal{N}(0,1)$$
$$\bar{X}_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$
$$\bar{X}_n - \mu \approx \mathcal{N}\left(0, \frac{\sigma^2}{n}\right)$$
$$\sqrt{n}(\bar{X}_n - \mu) \approx \mathcal{N}\left(0, \sigma^2\right)$$
$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \approx \mathcal{N}(0,1)$$

Continuity correction

$$\mathbb{P}\left[\bar{X}_n \le x\right] \approx \Phi\left(\frac{x + \frac{1}{2} - \mu}{\sigma/\sqrt{n}}\right)$$

$$\mathbb{P}\left[\bar{X}_n \ge x\right] \approx 1 - \Phi\left(\frac{x - \frac{1}{2} - \mu}{\sigma/\sqrt{n}}\right)$$

Delta method

$$Y_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \implies \varphi(Y_n) \approx \mathcal{N}\left(\varphi(\mu), (\varphi'(\mu))^2 \frac{\sigma^2}{n}\right)$$

11 Statistical Inference

Let $X_1, \cdots, X_n \overset{iid}{\sim} F$ if not otherwise noted.

11.1 Point Estimation

- Point estimator $\widehat{\theta}_n$ of θ is a RV: $\widehat{\theta}_n = g(X_1, \ldots, X_n)$
- $\mathsf{bias}(\widehat{\theta}_n) = \mathbb{E}\left[\widehat{\theta}_n\right] - \theta$
- Consistency: $\widehat{\theta}_n \xrightarrow{\text{P}} \theta$
- Sampling distribution: $F(\widehat{\theta}_n)$
- Standard error: $\mathsf{se}(\widehat{\theta}_n) = \sqrt{\mathbb{V}\left[\widehat{\theta}_n\right]}$