Probability Functions

General (*independent assumed)

$$\mathbb{E}[X] = \mu = \sum_x x p_X(x) = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$*\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

$$\text{Var}[X] = \mathbb{E}\left[(X-\mu)^2\right] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$*\,\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$$

$$\sigma_X \coloneqq \sqrt{\text{Var}[X]} = \sqrt{\mathbb{E}\left[(X-\mu)^2\right]}$$

$X \sim$ **Bernoulli**(p)

$$p_X(x) = \begin{cases} 1-p & \text{if } x = 0 \\ p & \text{if } x = 1 \end{cases}$$

$$\mathbb{E}[X] = p \qquad \text{Var}[X] = p(1-p)$$

$X \sim$ **Binomial**(n, p)

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mathbb{E}[X] = np \qquad \text{Var}[X] = np(1-p)$$

$X \sim$ **Geometric**(p)

$$p_X(k) = \mathbb{P}(X = k) = p(1-p)^{k-1}$$

$$\mathbb{E}[X] = \frac{1}{p} \qquad \text{Var}[X] = \frac{1-p}{p^2}$$

$X \sim$ **Poisson**(λ)

$$p(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, 3$$

$$\mathbb{E}[X] = \lambda \qquad \text{Var}[X] = \lambda$$

$X \sim$ **Exponential**(λ)

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

$$F_X(x) = 1 - e^{-\lambda x} \quad \mathbb{E}[X] = \frac{1}{\lambda} \quad \text{Var}[X] = \frac{1}{\lambda^2}$$

$X \sim$ **Uniform**(a, b) $dF_X(x) \Rightarrow f_X(x)$

$$F_X(x) = \begin{cases} 0, & \text{if } x < a \\ \dfrac{x-a}{b-a}, & \text{if } a \leq x \leq b \\ 1, & \text{if } b < x \end{cases}$$

$$\mathbb{E}[X] = \frac{a+b}{2} \qquad \text{Var}[X] = \frac{(b-a)^2}{12}$$

Normal Distribution

$$X \sim \mathcal{N}(\mu, \sigma^2) : f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\mathbb{E}[X] = \mu \qquad \text{Var}[X] = \sigma^2$$

Standard Normal Distribution

$$\mathcal{N}(0,1) : f_X(x) \coloneqq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\mathbb{E}[Z] = 0 \qquad \text{Var}[Z] = 1 \qquad Z = \frac{X-\mu}{\sigma}$$

Central Limit Theorem

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \; S_n = X_1 + \cdots + X_n$$

$$\lim_{n\to\infty} \mathbb{P}(Z_n \leq z) = \Phi(z)$$

Bayesian Statistic

Bayes' rule

$$f_{\Theta|X}(\theta|x) = \frac{f_\Theta(\theta) f_{X|\Theta}(x|\theta)}{f_X(x)} \propto f_{X|\Theta} f_\Theta(\theta)$$

$$\mathbb{P}(\theta|x_1, x_2) = \frac{\mathbb{P}(x_2|\theta, x_1)\mathbb{P}(\theta|x_1)}{\mathbb{P}(x_2|x_1)}$$

$$f_{X_1|\Theta}(x_1|\theta) \cdots f_{X_n|\Theta(x_n|\theta)} f_\Theta(\theta)$$

Beta Distribution

$$f_\Theta(\theta) = \begin{cases} \underbrace{\frac{1}{B(\alpha,\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}}_{\text{if } 0<\theta<1} \\ 0 \quad \text{otherwise} \end{cases}$$

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \Gamma(\alpha) = (\alpha - 1)!$$

$$B(\alpha, \beta) = \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!}.$$

$$\Theta \sim \text{Uniform}(0,1) = \text{Beta}(1,1)$$

$$\text{mode}[\theta] = \frac{\alpha - 1}{\alpha - 1 + \beta - 1} \quad \text{when } \alpha, \beta > 1$$

$$\text{Bernoulli: Beta}(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i)$$

$$\text{Posterior: } (\theta|h,t) \sim \text{Beta}(h+1, t+1)$$

Gamma Distribution

$$f_\Theta(\theta) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} & \text{for } \theta > 0 \\ 0 & \text{for } \theta \leq 0 \end{cases}$$

Poisson, prior Gamma(α, β), n = #exp

$$\text{Gamma}(\alpha + \sum_{i=1}^n x_i, \beta + n)$$

Exponential, prior Gamma(α, β),

$$\text{Gamma}(\alpha + n, \beta + \sum_{i=1}^n x_i)$$

OR $\alpha =$# trials + 1, $\beta =$ data sum + prior

Normal Distribution

prior: $\mathcal{N}(\mu, \sigma_0^2)$, posterior:

$$\mu' = \frac{\sigma^2\mu_0 + \sigma_0^2 \sum_{i=1}^n x_i}{\sigma^2 + n\sigma_0^2} \qquad \sigma'^2 = \frac{\sigma^2\sigma_o^2}{\sigma^2 + n\sigma_o^2}$$

General, prior $\mathcal{N}(\mu_0, \sigma_0^2)$, post $\mathcal{N}(\mu, \sigma^2)$

$\frac{\mu}{\sigma^2} = \frac{\mu_0}{\sigma_0^2} + \frac{x_1}{\sigma_1^2} + \cdot + \frac{x_n}{\sigma_n^2}$ $\frac{1}{\sigma^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma_1^2} + \cdot + \frac{1}{\sigma_n^2}$

Special, $\sigma_0^2 = \sigma^2 = 1$

$$\mu' = \frac{\mu_0 + \sum_{i=1}^n x_i}{1+n} \qquad \sigma'^2 = \frac{1}{1+n}$$

Prediction

$$\mathbb{P}(x^* \in [a,b] | X = x) = \int_-^+ \mathbb{P}(\cdot) f_{\Theta|X}(\theta|x) d\theta$$

$$\mathbb{P}(x^* \in [a,b]|\theta) = \int_a^b f_{X|\Theta}(x^*|\theta) dx^*.$$

Point estimation MAP - find post, cal

$$\theta_{\text{MAP}} = \arg\max_\theta f_{\Theta|X}(\theta|x)$$

prior Beta$(1,1)$, post Beta$(1+h, 1+t)$

$$\text{Beta: } \theta_{MAP} = \frac{\alpha-1}{\alpha-1+\beta-1} = \frac{h}{h+t}$$

prior $\mathcal{N}(\mu_0, 1)$, post $\mathcal{N}(\frac{\mu_0 + x_1 + \cdots + x_n}{n+1}, \frac{1}{n+1})$:

$$\text{Normal: } \theta_{MAP} = \frac{\mu_0 + x_1 + \cdots + x_n}{n+1}$$

Hypothesis Testing

$$\mathbb{P}(\hat\theta \neq \theta) = \mathbb{P}(\hat\theta = 1, \theta = 0) + \mathbb{P}(\hat\theta = 0, \theta = 1)$$

$$= \mathbb{P}(\hat\theta = 1|\theta = 0)\mathbb{P}(\theta = 0) + \ldots$$

Sampling Statistic

Sample mean: $\overline{X} = \dfrac{X_1 + \cdots + X_n}{n}$

Sample variance: $s^2 = \dfrac{\sum_{i=1}^n (X_i - \overline{X})^2}{n}$

$$\mathbb{E}[\overline{X}] = \mu \qquad \text{Var}[\overline{X}] = \frac{\sigma^2}{n} \qquad Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\overline{X} \sim \mathcal{N}\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \; \lim_{n\to\infty} \mathbb{P}\left(\overline{X} \leq \mu + t\frac{\sigma}{\sqrt{n}}\right)$$

$$\mathbb{E}[S^2] = \frac{n-1}{n}\sigma^2 \quad \mathbb{E}[\frac{n}{n-1}S^2] = \sigma^2(\text{unbiased})$$

Max Likelihood Estimation

Unbiased: $\mathbb{E}[\widehat{\Theta}_n] = \theta$

Asymptotically unbiased: $\lim_{n\to\infty} \mathbb{E}[\widehat{\Theta}_n] = \theta$

Consistent: $\widehat{\Theta}_n$ converges to θ

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \mathbb{P}(|\widehat{\Theta}_n - \theta| \geq \varepsilon) = 0$$

$$\hat\theta_n = \arg\max_\theta f_X(x_1, \cdots, x_n|\theta) \; \theta \text{ unknown}$$

$$\text{Bernoulli: } \theta_{\text{MLE}} = \frac{k}{n}$$

$$\frac{\partial f_X(x_1, \cdots, x_n|\theta)}{\partial\theta} = 0$$

$$\hat\theta = \arg\max_\theta \ln\left(f_X(x_1, \cdots, x_n|\theta)\right)$$

Find likelihood functions, ln, differentiate

$$\text{max likelihood} \begin{cases} \hat\mu = \dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i \\ \hat\sigma^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^n (X_i - \hat\mu)^2 \end{cases}$$

Unbiased estimator: $\hat\sigma^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (X_i - \hat\mu)^2$