Formula Sheet

Hoeffding's Inequality: [if $a \le x \le b$ - bounded RV]

$\varepsilon = \sqrt{\frac{(b-a)}{2n} \ln\left(\frac{2}{a}\right)}$ $\qquad I = (\bar{X}_n - \varepsilon, \bar{X}_n + \varepsilon)$ $\qquad \mathbb{P}(m \text{ is in } I) \geq 1-a$

→ The true mean is inside the interval, and will be correct 1-a% of times

Independent Events $P(AB) = P(A) \cdot P(B)$ $\qquad P(AB) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$

if independent: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$

Bayes Theorem: $P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_j P(B|A_j) \cdot P(A_j)}$

Law of Total Probability: $P(B) = \sum P(B|A_i) \cdot P(A_i)$

$E(X) = \int x\,dF(x) = \begin{cases} \sum_x x \cdot f(x) & \text{if discrete} \\ \int x\, f(x)\, dx & \text{if continuous} \end{cases}$

$E(a \cdot X) = aE(X)$ $\qquad E[X+Y] = E[X] + E[Y]$ $\qquad E[XY] = E[X] \cdot E[Y]$, if independent

$E[g(x)] = \int g(x) \cdot dF(x) = \begin{cases} \sum g(x) f(x) & \text{if discrete} \\ \int_{-\infty}^{+\infty} g(x) f(x)\, dx & \text{if continuous} \end{cases}$

Rule of Lazy Statistician

eg. $g(x) = x^2$, $f(x) = P(X = x)$ → $\sum 0^2 f(0) + 1^2 f(1) + \ldots$

eg. $X \sim \text{Uniform}(0,1)$ $\quad f(x) = \begin{cases} 0 & \text{otherwise} \\ 1, & 0 \le x \le 1 \end{cases}$

$g(x) = x^2 \quad E[g(x)] = \int_0^1 x^2 f(x)\,dx = \int_0^1 x^2\,dx$

Discrete RV

1. Bernoulli: $f(x) = p^x (1-p)^{1-x}$, $x \in \{0, 1\}$
2. Binomial: $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$
3. Geometric: $p(1-p)^{x-1}$
4. Poisson: $f(x) = e^{-\lambda} \frac{\lambda^x}{x!}$

Continuous RV

1. Uniform $f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$
2. Normal (Gaussian): $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$
3. Exponential $f(x) = \frac{1}{b} e^{-x/b}$
4. Gamma $f(x) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-x/b}$
5. Beta
6. Cauchy
7. χ^2

$L(a, b) = (a-b)^2$: Quadratic Loss $\quad R(g) = \mathbb{E}[(g(x) - Y)^2]$ - Least Square (cares about big errors)

$L(a, b) = |a-b|$: Absolute Loss $\quad R(g) = \mathbb{E}[|g(x) - Y|]$ - treats every error proportionally (same)

Convergence:

1) The Law of Large Numbers: converges in probability $X_n \xrightarrow{P} X$

2) The Central Limit Theorem: converges in Distribution (Normal) $X_n \rightsquigarrow X$

which patient features are reasonable to predict target "number of physician office visits"

emr ?	hldldiff	famihc
hosp ?	hortast	employed
poorhealth	midwest	privins
numchron	age	medicaid
	male	

$$f_{Y|X}(y, x) = \frac{\lambda^y}{y!} e^{-\lambda} \qquad (\lambda(x) = \exp(ax+b))$$

$$\log f_{Y|X}(y, x) = \log\left(\frac{\lambda^y}{y!} e^{-\lambda}\right) = \log(\lambda^y) + \log(e^{-\lambda}) = y\log(\lambda) - \lambda$$

using $\lambda(x) = \exp(ax+b)$:

$$= y \cdot \log(\exp(ax+b)) - \exp(ax+b) = y(ax+b) - \exp(ax+b)$$

to maximise the ~~log~~ Negative log-likelihood, we need to min. it.

$$L(a, b) = -\sum_{i=1}^{n} (y_i(ax_i+b) - \exp(ax_i+b))$$

$$\Rightarrow L(a, b) = \sum_{i=1}^{n}\left[\exp(ax_i+b) - y_i(ax_i+b)\right]$$