* Information Theory

- $I(p) \geq 0$, $I(1) = 0$, $I(p_1 \cdot p_2) = I(p_1) + I(p_2)$

  $I(p) = -\log_b P$

  · Information Content
  $\begin{cases} b=2 : \text{bits} \\ b=3 : \text{trits} \\ b=e : \text{nats} \\ b=10 : \text{Hartleys} \end{cases}$

- Entropy $H(X) = E[I(X)] = -\sum_{i=1}^{n} p(x_i) \cdot \log_b p(x_i)$

- Relative Entropy $D(P\|Q) = E_P\left[\log_b \frac{p(x)}{q(x)}\right] = \sum_x p(x) \cdot \log_b \frac{p(x)}{q(x)}$

- Joint Entropy $H(X,Y) = E[-\log_b P(X,Y)]$

  $= -\sum_x \sum_y P(X=x, Y=y) \log_b P(X=x, Y=y)$

  ① $H(X,Y) \geq \max(H(X), H(Y))$

  ② $H(X,Y) \leq H(X) + H(Y)$

  ③ $H(X,Y) = H(X) + H(Y)$ if, only if X, Y are independent

- Conditional Entropy $H(X|Y) = H(X,Y) - H(Y)$

  $= -\sum_x \sum_y P(X=x, Y=y) \log_b P(X=x | Y=y)$

  $= -\sum_x \sum_y P(X=x, Y=y) \log_b P(X=x, Y=y) + \sum_y P(Y=y) \log_b P(Y=y)$

- Mutual Information $I(X,Y) = H(X) - H(X|Y)$

  I↑ → X,Y 겹치는 것↑ $= H(Y) - H(Y|X)$

  I↓ → X,Y 〃 ↓ $= H(X) + H(Y) - H(X,Y)$

* Central limit theorem

$\bar{X}$ follows approximately normal distribution with $(\mu, \frac{\sigma}{\sqrt{n}})$ from X,

even if X does not follow normal distribution,

if $n > 30$

* $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$ 일 때, $Var(\bar{X}) = \sqrt{n \times \sigma^2}$

* Point Estimation

- Case 1 : $\bar{X}$, σ is given , $\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

- Case 2 : population follows Bernauli Distribution

  With given $\bar{X}$ without σ

  · Upper Bound : $\bar{X} \pm \frac{z_{\alpha/2}}{2\sqrt{n}}$

  · Estimation : $\bar{X} \pm z_{\alpha/2} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$

- Case 3 : $\bar{X}$ is given with unknown σ (distribution)

  $\bar{X} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$

* Hypothesis Testing

- z-test : the sample follows a normal distribution
  $n > 30$

  cf. normal dist. 이면 n<30 이라도 z-test

- t-test : $5 \leq n \leq 30$ 또는
  s가 unknown 일 때?

  $\boxed{T_{n-1} = \frac{\bar{X} - \mu}{\frac{S_n}{\sqrt{n}}}}$

  $\boxed{S_n = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}}$

  ① One sample T-test

  ② Two samples T-test

  $\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

  $T_{n_1+n_2-2} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}}$ — $\mu_1 = \mu_2$

  ③ Paired sample T-test

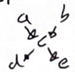
  i. $H_0 : \mu \quad \mu_0$
  $H_1 : \mu \quad \mu_0$

  ii. $\alpha = \cdots$, $z_\alpha = \cdots$

  iii. $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

  iv. accept / reject
  - accept H_1 if z is in tails.

* Bayesian Network



$p(a, b, c, d, e) = p(e|c) \cdot p(d|c) \cdot p(c|a,b) \cdot p(a) \cdot p(b)$

* Pearson Correlation

$$\rho_{X,Y} = \frac{1}{\sigma_X \sigma_Y} \sum (x-\mu_X)(y-\mu_Y) \cdot P(X=x, Y=y)$$

- X, Y independent $\rightarrow \rho_{X,Y} = 0$

  not vice versa

* Sample Correlation Coeff.

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

* Linear Regression

$Y = aX + b$

$a = \text{Corr} \times \frac{S_Y}{S_X}$ (s: standard deviation)

$b = M_Y - a \cdot M_X$

$$* \sigma^2 = \sum (x_i - \mu)^2 \cdot p_i$$

확률이 모두 같으면 $\frac{1}{n}$이 되므로 population에서 N으로 나누는 거다.

or $E(X^2) - (E(X))^2$

* Maximum Likelihood Estimation

- when p_1, p_2, p_3 is given,

  Find the highest probability for

  $$P(p = p_i \mid \text{sample condition})$$

ex. if a coin (might unfair) observed 13 Heads in 20 trials with estimated Prob. $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}$

① $p(p=\frac{1}{2} \mid H=13) = {}_{20}C_{13} \cdot (\frac{1}{2})^{13} (\frac{1}{2})^{7} = \cdots$

⋮

$p = \frac{2}{3}$ is MLE.