$P(A) \geq 0$

$P(\Omega) = 1$

$$P(\underset{A \cup B}{A \text{ or } B}) = P(A) + P(B) - P(\underset{A \cap B}{A \text{ and } B})$$

$$P(A) = P(\underset{A \cap B}{A \text{ and } B}) + P(\underset{A \cap B^c}{A \text{ and not } B})$$

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

If A and B are independent

$$P(A \text{ and } B) = P(A)\,P(B)$$

$${}_nP_k = \frac{n!}{(n-k)!} \qquad {}_nC_k = \binom{n}{k} = \frac{{}_nP_k}{k!} = \frac{n!}{k!(n-k)!}$$

- Schwarzenegger's veto message
- The birthday problem

$$P(A \text{ and } B) = P(A|B)\,P(B) \quad \left(\because P(A|B) = \frac{P(A \text{ and } B)}{P(B)}\right)$$

$$P(A) = P(A|B)\,P(B) + P(A|B^c)\,P(B^c)$$

$$\because P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \qquad P(A|B^c) = \frac{P(A \text{ and } B^c)}{P(B^c)}$$

- the twin problem
- The Florida registered voter list example

- Monty Hall problem

MC : Monty reveals door C

$P(A) = P(B) = P(C)$ (e.g. A: car behind door A)

$$P(A \mid MC) = \frac{P(A \text{ and } MC)}{P(MC)} = \frac{P(A)\,P(MC \mid A)}{P(MC)}$$

$$P(B \mid MC) = \frac{P(B \text{ and } MC)}{P(MC)} = \frac{P(B)\,P(MC \mid B)}{P(MC)}$$

$P(MC \mid A) = \frac{1}{2}$ (Monty could choose B or C)

$P(MC \mid B) = 1$ (Monty could choose only C)

Bernoulli distribution

Binomial distribution

Uniform distribution

Normal distribution

$X \sim N(\mu, \sigma^2)$

$Y = aX + b \Rightarrow Y \sim N(a\mu + b, (a\sigma)^2)$

$$Pr(k \le x \le k') = P(x \le k') - P(x \ge k)$$
$$= F(k') - F(k)$$

P : probability distribution function for $N(\mu, \sigma^2)$

F : Cumulative " "

Equalities for expectation & Variance

Expection & variance of distributions

Definitions of Unbiasedness and Consistency

Central limit theorem for $\bar{X}_n$ (Sample mean of size n)

Standard error

Hypothesis testing

- Tea tasting experiment example

Confidence interval : definition & calculation

Linear regression : regression table
standard error of $\hat{\beta}$ (don't need to calculate)
Confidence interval of $\hat{\beta}$
(need to calculate given s.e.)
hypothesis testing for $\hat{\beta}$
standard error & confidence interval for $\hat{Y}$ (predicted outcome)

Interpretation for Average Treatment Effect

Assumptions of Linear regression models

- Omitted variable bias

- RCTs
- Observational studies

population X ; mean μ, var $V(X)$

sample mean $\bar{X}_n$: $E(\bar{X}_n)=\mu$, $V(\bar{X}_n)=V(X)/n$

$H0: \mu=\mu_0$

test statistic

$$z=\frac{\bar{X}_n-\mu}{\sqrt{V(\bar{X}_n)}}=\frac{\bar{X}_n-\mu}{\sqrt{V(X)/n}}$$

Hypothesis testing:

we assume H_0 is correct

⇓

μ_0

Use z or t to conduct testing

→ critical values

→ t-test

If $V(X)$ is unknown

$V(\bar{X}_n)\approx S^2/n$ (standard error: $\sqrt{S^2/n}$)

S^2 (sample variance) $=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{X}_n)^2$

$$t=\frac{\bar{X}_n-\mu}{\sqrt{S^2/n}}$$

Use t-distribution (degree of freedom)

↙

$n-1$

for one-sample test

2 sample test

population X_1 : mean μ_1, Var $V(X_1)$

population X_2 : mean μ_2, Var $V(X_2)$

$H_0: \mu_1=\mu_2 \Rightarrow \mu_1-\mu_2=0$

$$z=\frac{(\bar{X}_{1n_1}-\bar{X}_{2n_2})-0}{\sqrt{V(\bar{X}_{1n_1}-\bar{X}_{2n_2})}}$$

$V(\bar{X}_{1n_1}-\bar{X}_{2n_2})$

$=V(\bar{X}_{1n_1})+(-1)^2V(\bar{X}_{2n_2})$

$=V(X_1)/n_1+V(X_2)/n_2$

If $V(X_1)$ and $V(X_2)$ not known,

$$t=\frac{(\bar{X}_{1n_1}-\bar{X}_{2n_2})-0}{\sqrt{S_1^2/n_1+S_2^2/n_2}}$$

For proportions, recall Bernoulli RV

population X : mean p, Var $p(1-p)$

no need for additional information than mean p

$H_0: P=P_0$

$$z=\frac{\bar{X}_n-\mu}{\sqrt{V(\bar{X}_n)}}=\frac{\bar{X}_n-p}{\sqrt{p(1-p)/n}}$$

2 sample test

$$z=\frac{(\hat{p}_1-\hat{p}_2)-0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

$\hat{p}_1$: sample 1 mean

$\hat{p}_2$: sample 2 mean