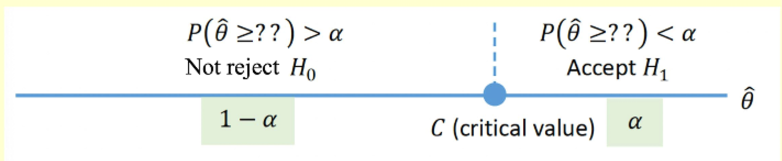
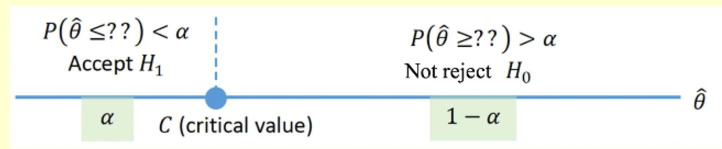
Topic 3

- **Martingale**
  - definition
    - **Definition 8.3**
      An adapted sequence $(X_n, \mathcal{F}_n)$ of real random variables is called a ***martingale*** if for all $n \in \mathbb{N}$ it holds that $E|X_n| < \infty$ and
      $$E(X_{n+1}|\mathcal{F}_n) = X_n \ a.s.,$$
      and a ***submartingale*** if for all $n \in \mathbb{N}$ it holds that $E|X_n| < \infty$ and
      $$E(X_{n+1}|\mathcal{F}_n) \geq X_n \ a.s.,$$
      and a ***supermartingale*** if $(-X_n, \mathcal{F}_n)$ **is a submartingale**.
  - Stopping times
    - definition
      - **Definition 9 (Discrete time)**
        A ***stopping time*** is a random variable $\tau: \Omega \to \mathbb{N} \cup \{\infty\}$ such that
        $$(\tau = n) \in \mathcal{F}_n$$
        for all $n \in \mathbb{N}$. If $\tau < \infty$ we say that τ is a finite stopping time.
      - We have an equivalent definition of a stopping time:
        A random variable $\tau: \Omega \to \mathbb{N} \cup \{\infty\}$ is a stopping time if and only if $(\tau \leq n) \in \mathcal{F}_n$ for all $n \in \mathbb{N}$.
    - property
      - Let $X_1, X_2, \ldots$ be a sequence of random variables, all defined on a common sample space Ω. Let τ be a nonnegative integer-valued random variable, also defined on Ω. We call τ a *stopping time adapted to the sequence* $\{X_n\}$ if $P(\tau < \infty) = 1$, and if for each $n \geq 1$, if $I_{\{\tau \leq n\}}$ is a function of only $X_1, X_2, \ldots X_n$.

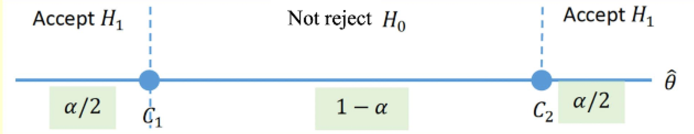
- **Random sampling**
  - point estimation
    - good estimatior
      - unbiased: $E[\hat{\theta}] = \theta$; Asymptotically unbiased estimator: $\lim_{n\to\infty} E[\hat{\theta}] = \theta$
      - consistency: $\lim_{n\to\infty} P(|\hat{\theta} - \theta| > \epsilon) = 0$ (像是$\hat{\theta}$依概率P收敛到θ，但unbiased 又不是a.s.，中间还添加了一个期望，两个性质之间没有关系)
      - MSE: $MSE = E[(\hat{\theta} - \theta)^2] = VAR(\hat{\theta}) + bias(\hat{\theta})^2$
        - $bias = E[\hat{\theta}] - \theta$
        - UMVUE：MSE最小的那个统计量
          - how to find a UMVUE?
            - T是一个统计量，由X来计算的统计量
            - sufficiency: 已知T的情况下，X的概率分布和真实的参数独立<--只要知道T，那么知不知道真实参数对我来说没有关系了，所以称为是sufficient
              - theorem
                - **Theorem (Neyman-Fisher Factorization theorem: Necessary and sufficient conditions for sufficiency)**
                  A statistic $T = T(X_1, X_2, \ldots, X_n)$ is sufficient for parameter θ if and only if the joint density function (probability density function or probability mass function) $f(x_1, \ldots, x_n|\theta)$ can be expressed as
                  $$f(x_1, \ldots, x_n|\theta) = g(T(x_1, x_2, \ldots, x_n), \theta) \times h(x_1, \ldots, x_n)$$
                  where $g(\cdot)$ and $h(\cdot)$ are some nonnegative functions.
              - property
                - **Theorem (Rao-Blackwell Theorem)**
                  We let $T = T(X_1, X_2, \ldots, X_n)$ be a sufficient statistics for parameter θ and $\hat{\theta}_n$ be an unbiased estimator for θ. We define an estimator $\tilde{\theta}$ as $\tilde{\theta} = \mathbb{E}[\hat{\theta}_n|T]$.
                  Then $\tilde{\theta}$ is unbiased estimator and $\mathbb{E}\left[(\tilde{\theta} - \theta)^2\right] \leq \mathbb{E}[(\hat{\theta}_n - \theta)^2]$. ① ⇒ $\tilde{\theta}$ unbiased
            - completeness
              - definition
                - We say a statistic $T = T(X_1, X_2, \ldots, X_n)$ is complete if and only if the following condition holds: If $\mathbb{E}_\theta(g(T)) = 0$ for all θ, then $g(T) = 0$.
              - property
                - **Theorem**
                  We let $T = T(X_1, X_2, \ldots, X_n)$ be a sufficient and complete statistic for a parameter θ and $\hat{\theta}_n$ be any unbiased estimator, then the estimator
                  $$\tilde{\theta} = \mathbb{E}[\hat{\theta}_n|T]$$
                  is the unique UMVUE.
                  (*In other words, the estimator generated from Rao-Blackwell theorem is the desired UMVUE.)
            - 既然有这么好的性质，那怎么判断一个T是充分完备统计量呢?
              - property
                - **Definition 6 (k-parameter exponential families)**
                  A family of distribution $\{P_\theta\}$ (where $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ is k-dimension parameter) is called k -parameter exponential family if the probability density function (or probability mass function) can be expressed as
                  $$f(X|\theta) = \mathbf{1}_{\{X \in A\}} e^{\sum_{i=1}^{k} c_i(\theta) T_i(X) + d(\theta) + S(X)}$$
                  where $c_i(\cdot),\ d(\cdot),\ S(\cdot)\ and\ T_i(\cdot)$ are some functions.
                  Here, A denotes the collection of possible values of the random variable (X) and it must be independent of θ.
                  Let $X_1, \ldots, X_n$ be an $i.i.d$ sample from a k-parameter exponential family, then ($S_1 = \sum_{i=1}^{n} T_1(X_i)$, …, $S_k = \sum_{i=1}^{n} T_k(X_i)$) is a set of complete and sufficient statistic for $(\theta_1, \ldots, \theta_k)$.
  - interval estimation
    - 如果仅仅是点估计的话，可能会导致样本外数据的拟合效果很差，所以我们引入了区间估计
    - z检验，t检验
    - two population mean estimation
      - Theorem
        If $X_1, X_2, \ldots, X_n \sim N(\mu_X, \sigma^2)$ and $Y_1, Y_2, \ldots, Y_m \sim N(\mu_Y, \sigma^2)$ are independent random samples, then a $(1-\alpha)100\%$ confidence interval for $\mu_X - \mu_Y$, the difference in the population means is:
        $$(\bar{X} - \bar{Y}) \pm (t_{\frac{\alpha}{2}, n+m-2})\, S_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$
        where $S_p^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$, the "pooled sample variance",
        S_p^2 is an unbiased estimator of the common variance σ^2.
    - hypothesis testing
      - one tail test
        - $H_0 :\leq$
          - **Case 1: One tail test ($H_0: \theta = \theta_0, H_1: \theta > \theta_0$)**

        - $H_0 :\geq$
          - **Case 2: One tail test ($H_0: \theta = \theta_0, H_1: \theta < \theta_0$)**

      - Two tail test
        - **Case 3: Two tail test ($H_0: \theta = \theta_0, H_1: \theta \neq \theta_0$)**

    - p value approach
      - p是现在实际情况的概率，如果H_0:<=, 那么p是右边的概率，p小于α，说明X落在拒绝域，拒绝H_0