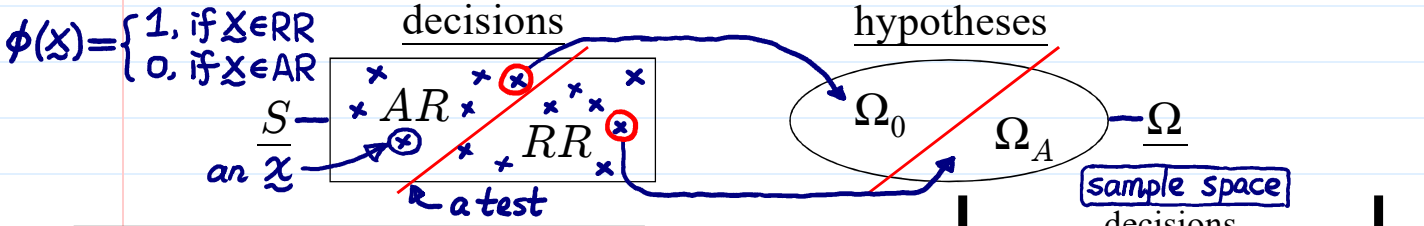
• Neyman-Pearson paradigm --- concept and procedure of hypothesis testing

Definition 7.2 (test, rejection region, acceptance region, test statistic, TBp. 331)

- A **test** is a rule defined on sample space of data (denoted as S) that specifies:
  1. for which sample values $\mathbf{x} \in S$, the decision is made to accept H_0
  2. for which sample values $\mathbf{x} \in S$, reject H_0 in favor of H_A
- The subset of the sample space for which H_0 will be rejected is called the **rejection region**, denoted as RR.
- The complement of the rejection region is called the **acceptance region**, denoted as $AR = S \setminus RR = RR^c$.



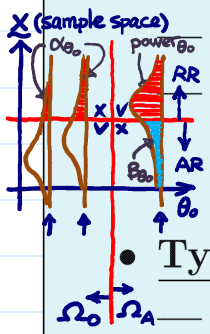
Note. Two types of errors may be incurred in applying this paradigm

		decisions: accept H_0	decisions: reject H_0
hypotheses	H_0 is true	Bingo!	Type I error
	H_A is true	Type II error	Bingo!

Definition 7.3 (type I error, significance level, type II error, power, TBp. 331)

Let Θ_0 denote the true value of parameters.

- **Type I error**: reject H_0 when H_0 is true, i.e., $\mathbf{x} \in RR$ when $\Theta_0 \in \Omega_0$.
  - α_{Θ_0} = probability of Type I error = $P(RR|\Theta_0)$, where $\Theta_0 \in \Omega_0$.
  - **Significance level** α: the maximum (or the least upper bound) of the Type I error probability, i.e., $\alpha = \max_{\Theta \in \Omega_0} \alpha_\Theta$ or $\alpha = \sup_{\Theta \in \Omega_0} \alpha_\Theta$.
- **Type II error**: accept H_0 when H_A is true, i.e., $\mathbf{x} \in AR$ when $\Theta_0 \in \Omega_A$
  - β_{Θ_0} = probability of Type II error = $P(AR|\Theta_0)$, where $\Theta_0 \in \Omega_A$
  - For $\Theta_0 \in \Omega_A$, $\mathbf{power}_{\Theta_0} = 1 - \beta_{\Theta_0} = P(RR|\Theta_0)$ = probability of rejecting H_0 when H_A is true



Question 7.4 (dilemma between type I and type II errors)

An ideal test would have $\alpha_\Theta = 0$, for $\Theta \in \Omega_0$ and $\beta_\Theta = 0$, for $\Theta \in \Omega_A$; but this can be achieved only in trivial cases. (Example?) In general, when α decreases, β increases, and vice versa.

Q: Can we find test s.t. α is minimized, and β is minimized?

smaller α, larger β — S | AR | RR — larger α, smaller β

A solution to the dilemma

- Strong evidence H_A is true → reject H_0
- Otherwise → accept H_0

The Neyman-Pearson approach imposes an asymmetry between H_0 and H_A: ↳ answer to question 7.3 in LNp.2

1. the significance level α is fixed in advance, usually at a rather small number (e.g. 0.1, 0.05, and 0.01 are commonly used), and (α: Type I error; small number → protect H_0)
2. then try to construct a test yielding a small value of β_Θ for $\Theta \in \Omega_A$. (β_Θ: Type II error)

$\arg\min_{\text{tests}} \beta_\theta$ subject to tests with $\sup_{\theta\in\Omega_0} \alpha_\theta \le \alpha$ — level-α tests

Example 7.2 (cont. Ex. 7.1 item 1 (LNp.3), TBp.330-331)

- Testing the value of the parameter p of a Bernoulli distribution with 10 trials. Let X be the number of successes, then $X \sim B(10, p)$ ← statistical modeling

Data → $\Omega = \{B(10,0.5), B(10,0.7)\}$

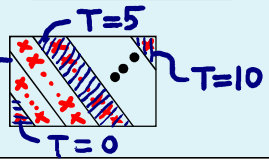
$$H_0 : p = 0.5 \quad \text{vs.} \quad H_A : p = 0.7$$

- *Different observations* of X give *different support* on H_A

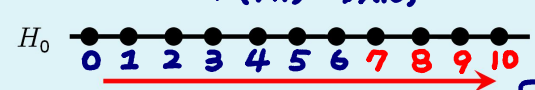
$X_1, \cdots, X_{10} \sim B(1, p)$ → $X = T(X_1, \ldots, X_{10})$

$T = X = \sum_{i=1}^{10} X_i$

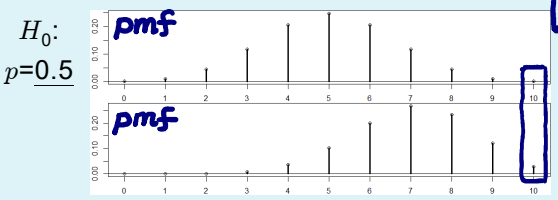
$S = \{(x_1, \cdots, x_{10}) \mid x_i \in \{0,1\}, i = 1, \cdots, 10\}$



$T(X_1, \cdots, X_{10})$



more support on H_A ← more extreme ← How to quantify it?



  - In this case, larger observations of X give more support on H_A.
  - Rejection region should consist of large values of X.
  - The observations that give more support on H_A are called "*more extreme*" observations

↳ most extreme: X=10, 2nd most extreme: X=9, ...

- If the rejection region is $\{7, 8, 9, 10\}$, then a test is formed. ($X \sim B(10, 0.5)$ under H_0)

$$\alpha = P(X \in \{7,8,9,10\} \mid p = 0.5) = 1 - P(X \le 6 \mid p = 0.5) = 0.18.$$

RR (H_0) ; AR (H_0)

- *Set the significance level as* 0.18.
  - The probability of Type II error (i.e., β) is ($X \sim B(10, 0.7)$ under H_A)

$$P(X \notin \{7,8,9,10\} \mid p = 0.7) = P(X \le 6 \mid p = 0.7) = 0.35,$$

AR (H_A) ; AR (H_A)

  - and the power (i.e., $1 - \beta$) is

$$P(X \in \{7,8,9,10\} \mid p = 0.7) = 1 - P(X \le 6 \mid p = 0.7) = 0.65.$$

RR (H_A) ; AR (H_A)

- Suppose that we are interested in testing

$\Omega = \{p \mid p \in [0,1]\}$ ← (≤)

$$H_0 : p = 0.5 \quad \text{vs.} \quad H_A : p > 0.5$$ ← $\Omega = \{p \mid p \in [0.5, 1]\}$

and the rejection region is still set to be $\{7, 8, 9, 10\}$:

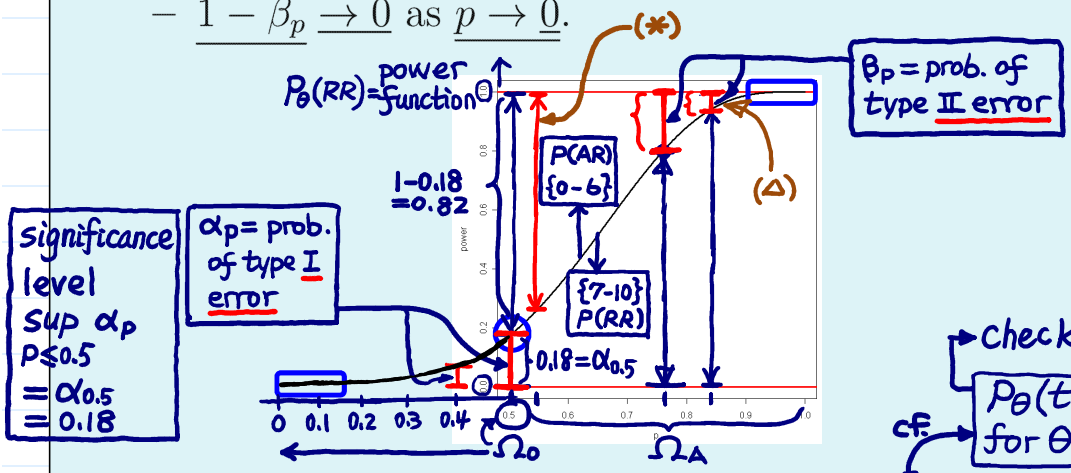
– Note that β_p is a function of p:

$$\beta_p = P(X \notin \{7,8,9,10\} \mid p\,).$$

power function $\equiv 1-\beta_p$.

– $1-\beta_p \to 1$ as $p \to 1$, and $1-\beta_p \to \alpha$ as $p \to 0.5$.

– $1-\beta_p \to 0$ as $p \to 0$.



Question 7.5

1. Which hypothesis get more *protection*? What kind of protection?
2. Can the protection make $P(\text{Type I error})$ always smaller than $P(\text{Type II error})$, i.e., $\alpha < \beta_\Theta$ for any $\Theta \in \Omega_A$?