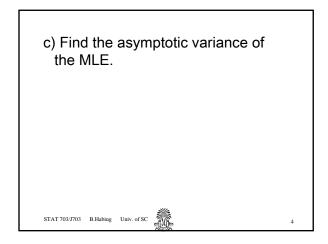


<u>Ch.8#5</u> Suppose that X follows a geometric distribution, *P*[*X* = *k*] = *p*(1−*p*)<sup>*k*−1</sup>
And assume an i.i.d. sample of size *n*.
b) Find the mle of *p*.





	0.400 0		
<u>Ch.8#6</u> Consider the data			
<u># Hops Freq</u>		<u># Hop</u>	<u>s Freq</u>
1	48	7	4
2	31	8	2
3	20	9	1
4	9	10	1
5	6	11	2
6	5	12	1
STAT 703	J703 B.Habing Univ. of S	c Min	5



a) Fit a geometric distribution

b) Find an approximate 95% confidence interval for p.

6

STAT 703/J703 B.Habing Univ. of SC

 $\frac{Ch8\#44c}{variables} \text{ Let } X_1, \dots X_n \text{ be i.i.d random}$  variables with density function

 $f(x \mid \theta) = (\theta + 1)x^{\theta}, 0 \le x \le 1$ 

Find the asymptotic variance of the MLE.

STAT 703/J703 B.Habing Univ. of SC

<u>Statistical Inference – Confidence</u> Intervals and Tests of Hypotheses

Test of hypothesis – general method to distinguish between 2 (or more) probability distributions (or models), based on a sample X<sub>1</sub>, ..., X<sub>n</sub> assumed to come from one of them.

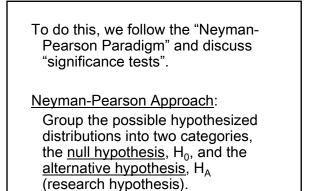
8

STAT 703/J703 B.Habing Univ. of SC

In particular, based on  $X_1, ..., X_n$ , decide whether  $f_1(x)$  or  $f_2(x)$  is the pdf (or population) from which the sample came.

<u>More specifically</u>, suppose we think the sample is from a normal population with mean either  $\mu$ =5 or  $\mu$ =10 with variance 4. (Or, more generally,  $\mu$ = $\mu$ <sub>1</sub> vs.  $\mu$ = $\mu$ <sub>2</sub>).

STAT 703/J703 B.Habing Univ. of SC



10

11

12

STAT 703/J703 B.Habing Univ. of SC

E.g. Observe X<sub>1</sub>, ..., X<sub>n</sub>. Either H<sub>0</sub>: N( $\mu_1$ ,  $\sigma^2$ ) or H<sub>A</sub>: N( $\mu_2$ ,  $\sigma^2$ ) or H<sub>0</sub>:  $\mu$ = $\mu_1$  vs. H<sub>A</sub>:  $\mu$ = $\mu_2$ ,  $\mu$  is the mean of N( $\mu$ ,  $\sigma^2$ ). Here, if  $\sigma^2$  is known, each of these hypotheses <u>completely</u> specifies

the distribution or population. So, H<sub>0</sub> and H<sub>A</sub> are called <u>simple</u> <u>hypotheses</u>.

STAT 703/J703 B.Habing Univ. of SC

If  $H_0$ :  $\mu$ =0 vs.  $H_A$ :  $\mu$  > 0 ( $\mu$  in N( $\mu$ ,  $\sigma^2$ ), with  $\sigma^2$  known).

Then H<sub>0</sub> is simple and H<sub>A</sub> is a <u>composite hypothesis</u>, i.e. several normal distributions would satisfy it.

 $H_A$  is also referred to as a <u>one-sided</u> hypothesis.

If H<sub>0</sub>:  $\mu$ =0 vs. H<sub>A</sub>:  $\mu$ ≠0 ( $\mu$  in N( $\mu$ ,  $\sigma$ <sup>2</sup>),  $\sigma$ <sup>2</sup> known), then H<sub>A</sub> is a <u>two-sided</u> (composite) hypothesis.

Next, we set up the framework for "testing"  $H_0$  and  $H_A$  based on the sample.

13

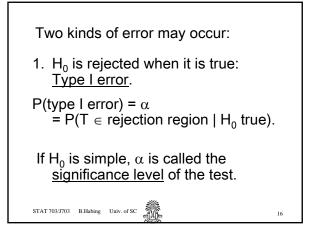
15

STAT 703/J703 B.Habing Univ. of SC

**9.2 Neyman-Pearson Paradigm** Let  $\underline{X} = (X_1, ..., X_n)$  denote a sample from population  $f(X|\theta)$ . Decide on  $H_0$  vs.  $H_A$  based on the sample. A decision on whether or not to reject  $H_0$  in favor of  $H_A$  is made on the basis of a statistic  $T=T(\underline{X})=T(X_1, ..., X_n)$ .

The set of values of T for which  $H_0$  is accepted is called the <u>acceptance region</u> and the set of values of T for which  $H_0$  is rejected is the <u>rejection region of the test</u>.

STAT 703/J703 B.Habing Univ. of SC



2. H<sub>0</sub> is accepted when it is false: <u>Type II error</u>.
P(type II error) = β
= P(T in acceptance region | H<sub>0</sub> false)
If H<sub>A</sub> is composite, β depends on which member of H<sub>A</sub> is the true pdf.

STAT 703/J703 B.Habing Univ. of SC

Power of the test=P(H₀ is rejected<br/>when false)= 1 - P(H₀ is accepted | H₀ false)= 1 - β.Ideally, we would want α=β=0, but<br/>this not possible since the decision<br/>is based on data.

17