MATH 425 11/4/2011

Note Title 11/4/2011

RECALL TEST 2 WEDNESDAY

DISCRETE RANDOM VARIABLES + GENERALITIES ON CONTINUOUS RANDOM VARIABLES

+ UNIFORM DISTRIBUTION

REVIEW MONDAY

STATISTICS VS. PROBABILITY

↗ TESTING HYPOTHESES.

↗ CALCULATING PROBABILITIES AND GENERALIZATIONS ⇝ STRATEGY

A TYPICAL EXAMPLE OF STATISTICS:

A COMPANY MAKES A NEW MEDICINE.

THIS WILL BE TESTED ON A SAMPLE OF PATIENTS.

EVALUATION OF OUTCOMES, vs. DISTRIBUTION OF OUTCOMES WITHOUT MEDICATION.

NULL HYPOTHESIS

SUPPOSE WE CAN TEST THE PROBABILITY p OF THE MEASURED OUTCOME (WITH DRUG) UNDER THE NULL HYPOTHESIS.

χ^2 TEST

VERY GENERAL RULE:

STATISTICALLY SIGNIFICANT MEANS $p < 0.05$

IF THIS HAPPENS, WE SAY THE DRUG HAD AN EFFECT. (REJECT THE NULL HYPOTHESIS).

IF $p \geq 0.05$, WE SHOULD NOT SAY ANYTHING.

WHY? SUPPOSE WE FOLLOW THIS SCIENTIFIC METHOD.

BAYESIAN ANALYSIS:

THE SCIENTIST
↓
CHANCE OF ME BEING WRONG

p **NULL HYPOTHESIS CORRECT** < 0.05

$1-p$ **NULL HYPOTHESIS FALSE** 0

Probability of being wrong is $< \underbrace{p}_{\leq 1} \cdot 0.05 < 0.05$.

THE CHI SQUARED TEST

χ^2

A VERY SIMPLE TEST.

PEARSON TEST: DRAW N BALLS (WITH RETURN)

k COLORS, PROBABILITY OF

i'TH COLOR IS p_i.

$$p_1 + p_2 + \cdots + p_k = 1.$$

NULL HYPOTHESIS – NUMBERS OF INDIVIDUAL BALLS FOLLOW THE DISTRIBUTION FROM EQUAL OUTCOMES. (APPROXIMATELY NORMAL)

THE χ^2 TEST: SUPPOSE I GOT a_i BALLS OF COLOR i.

$$\chi^2 = \frac{(a_1 - p_1 N)^2}{p_1 N} + \frac{(a_2 - p_2 N)^2}{p_2 N} + \cdots + \frac{(a_k - p_k N)^2}{p_k N}$$

degrees of freedom: $k-1$.

Look at a χ^2 table with $k-1$ degrees of freedom, how likely it is? Compare with 0.05

(χ^2 with l degrees of freedom: sum of l independent standard Gaussian variables).

EXAMPLE: SUPPOSE I THROW A PAIR OF COINS 20 TIMES.

THE NUMBER OF TIMES I GOT:

p			
$1/4$	2 HEADS	2	a_1
$1/2$	1 HEAD	16	a_2
$1/4$	0 HEADS	2	a_3

$$\chi^2 = \frac{(2-5)^2}{5} + \frac{(16-10)^2}{10} + \frac{(2-5)^2}{5} = \frac{9}{5} + \frac{36}{10} + \frac{9}{5} =$$

$$= \frac{72}{10} = 7.2$$

2 degrees of freedom:

p	0.5	0.1	0.05
χ^2	1.386	4.605	5.991

WE **CAN** DISPROVE THIS IS RANDOM

(IT **IS** STATISTICALLY SIGNIFICANT).

IN THE CASE OF 1 DEGREE OF FREEDOM, ONE CAN UNDERSTAND WHY χ^2 IS THE SQUARE OF A GAUSSIAN:

$$p = p_1 \qquad p_2 = 1 - p \qquad a = a_1$$

$$\chi^2 = \frac{(a - pN)^2}{pN} + \frac{(N - a - (1-p)N)^2}{(1-p)N} =$$

$pN = N_1$

$$= (a - pN)^2 \left(\frac{1}{pN} + \frac{1}{(1-p)N} \right) =$$

$$= \frac{(a - pN)^2}{Np(1-p)} \sim$$ the square of a standard Gaussian!

NO HW TODAY, REVIEW MONDAY