Hypothesis testing

Hypothesis: A statement framed in terms of restritions on the statistical model.

* there is a reason to doubt the hypothesis if the outcoms we observed have low probability when the hypothesis is true.

Ex: Coin tossed →
claim $P(\text{tails}) = 0.8$ when tossed
Tossed 10 tims and lands tails up 4 tims.

Binomial $^{10}C_4 (0.8)^4 (0.2)^6 = 0.005$
when the claim is true. We can thus take the outcoms of the experiment as evidence against the claim.

On the other hand, 7 tails in the 10 tosses
$^{10}C_7 (0.8)^7 (0.2)^3 = 0.201$
when the claim is true, ~~and~~ we ~~can~~ have little reason to doubt the doubt the claim.

Statistical Hypothesis:

A statistical Hypothesis is some statement or assertion about a population or equivalently about the probability distribution charaterising a population which we want to verify on the basis of information available from the sample.

Test of Statistical Hypothesis:

A test of a statistical hypothesis is a two action decision problem after the experimental sample values have been obtained, the two actions being the acceptance or rejection of the hypothesis under consideration.

A statistical test procedure consists of

1. A statistical model
2. A null and alternative hypothesis expressed in terms of the statistical model
3. A function of the data, $T(x_1, x_2 \ldots x_n)$, the test statistics
4. A decision rule:

Reject H_0 (Null Hypothesis) in favor of H_1 for $T \in C$ (critical region) and fail to reject H_0 for $T \notin C$.

Null and Alternative Hypothesis :

If $x_1, x_2 \cdots x_n$ is a random sample of size n from a normal population with mean μ and variance σ^2, then the hypothesis

$$H_0 : \mu = \mu_0, \ \sigma^2 = \sigma_0^2$$

is a simple Hypothesis
(Null Hypothesis)

Where as each of the following hypothesis are Composite hypothesis ~~are~~ (Alternative Hypothesis).

i) $\mu = \mu_0$
ii) $\sigma^2 = \sigma_0^2$
iii) $\mu < \mu_0, \ \sigma^2 = \sigma_0^2$
iv) $\mu > \mu_0, \ \sigma^2 = \sigma_0^2$
v) $\mu = \mu_0, \ \sigma^2 < \sigma_0^2$
vi) $\mu = \mu_0, \ \sigma^2 > \sigma_0^2$
vii) $\mu < \mu_0, \ \sigma^2 > \sigma_0^2$

Null Hypothesis : ~~This~~ refers to
Any Hypothesis we wish to test and denoted by H_0.
The rejection of H_0 leads to the acceptance of an alternative hypothesis denoted by H_1.

- Reject H_0: in favor of H_1 because of sufficient evidence in the data

- Fail to reject H_0: because of insufficient evidence in the data.

$H_0 : P = 0.10$ → (Probability of defective)
$H_1 : P > 0.10$

Now 12 defective items 100 does not proof a $P = 0.10$ so the conclusion " fail to reject "
However, 20 defective out of 100 then 'reject H_0' in favour of $H_1 : P > 0.10$.

Testing a statistical Hypothesis:

A certain type of cold vaccine is known to be only 25% effective after a period of 2 years.

P = $\boxed{\text{20 people choosen}}$

If more than 8 people surpass 2 years without contracting the virus, then the new vaccine will be considered superior than previous one.

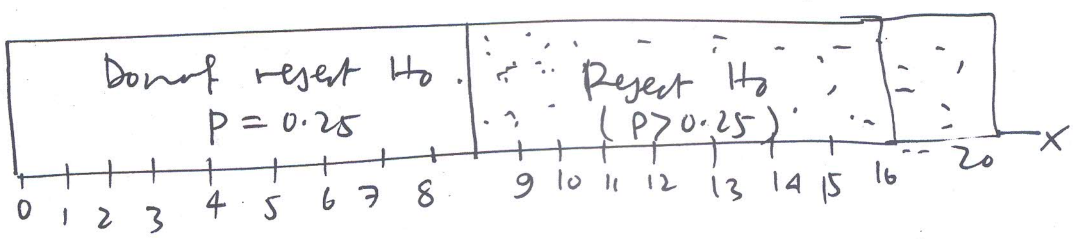
$\frac{8}{20} = 0.4$

$H_0: P = 0.25$

$H_1: P > 0.25$

$\boxed{\frac{8}{20} = \frac{4}{10} = 0.4}$

Test statistics:



The critical value is 8

if $x > 8 \rightarrow$ reject H_0 in favor of H_1

if $x \leq 8 \rightarrow$ we fail to reject H_0.

"The new vaccine may be no better than the one now in use"

~~$x > 8 \rightarrow$ reject H_0~~

Two types of Errors:

The decision to accept or reject the null hypothesis H_0 is made on the basis of the information supplied by the observed sample observations. the conclusion drawn on the basis of a particular sample may not always be true in respect of the population. The four possible situations that arise in any test procedure are given in the following Table

Double dichotomy (to oppose parts)/duality relating to decision and hypothesis:

		Decision From Sample	
		Reject H_0	Accept H_0
True Statement	~~H_0 True~~ H_0 is True	Wrong (Type I error)	correct
	H_0 is False (H_1 True)	Correct	Wrong Type II Error

Errors of Type I and Type II:

The error of rejecting H_0 (accepting H_1) when H_0 is true is called Type I error, and the error of accepting H_0 when H_0 is false (H_1 is true) is called type II error. The probabilites of type I and type II errors are denoted by α and β respectively.

α = Probability of rejecting H_0 when H_0 is true

β = Probability of accepting H_0 when H_0 is false.

* In quality control : $\alpha \rightarrow$ producer's risk
  $\beta \rightarrow$ Consumer's risk.

$$H_0: p = 1/4$$
$$H_1: p = 1/2$$

$$\begin{aligned}
\alpha &= P(\text{type I error}) \\
&= P(X > 8 \text{ when } p = 1/4) \\
&= \sum_{x=9}^{20} b(x; 20; 1/4) \\
&= 1 - \sum_{x=0}^{8} b(x: 20, 1/4) \\
&= 1 - 0.9591 \\
&= 0.0409
\end{aligned}$$

$$\begin{aligned}
\beta &= P(\text{type II error}) \\
&= P(X \leq 8 \text{ when } p = 1/2) \\
&= \sum_{x=0}^{8} b(x, 20, 1/2) \\
&= 0.2571
\end{aligned}$$

$$1 - \beta = 1 - 0.2571$$

Level of significance: α,

The probability of type 1 error, is known as the level of significance of the test. It is also known as the size of the critical region.

Power of the test: $1-\beta$ is ~~called~~ called the power function of the test hypothesis H_0 against the alternative hypothesis H_1. The value of the power function at a parameter point is called the power of the test at that point.

Let P be the probability that a coin will fall head in a single toss in order to test $H_0: P = 1/2$ against $H_1: P = 3/4$. The coin is tossed 5 times and H_0 is rejected if more than 3 heads are obtained. Find the probability of type I error and power of the test.

Sol^n. Here $H_0: P = 1/2$

and $H_1: P = 3/4$

If the random variable x denotes the number of heads in n tosses of a coin then

$X \sim B(n, p)$ so that

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \binom{5}{x} p^x (1-p)^{5-x}$$

$$\alpha = P(\text{type I error})$$

$$= P(X \geq 4 \text{ when } P = 1/2)$$

$$= P(X = 4 \mid P = 1/2) + P(X = 4 \mid P = 1/2)$$

$$= \binom{5}{4} (1/2)^4 (1/2)^{5-4} + \binom{5}{5} (1/2)^5$$

$$= \frac{3}{16}$$

$$\beta = \text{Probability (of Type II error)}$$

$$= 1 - [P(X = 4 \mid P = 3/4) + P(X = 4 \mid P = 3/4)$$

$$= 1 - \left[\binom{5}{4} (3/4)^4 (1/4) + \binom{5}{5} (3/4)^5\right]$$

$$= 1 - \frac{81}{218}$$

$$= 47/128$$

$\therefore$ Power of the test is

$$1 - \beta = \frac{18}{128} \quad //$$

Electronic Parts

(11) An ~~fabric~~ manufacturer believes that the proportion of orders for raw materials arriving late is $p=0.6$. If a random sample of 10 orders shows that 3 or fewer arrived late, the hypothesis that $p=0.6$ should be rejected in favor of the alternative $p<0.6$. use the binomial distribution

10.6

P-361

a) Find the probability of committing a type-I error if $p=0.6$

b) Find the probability of committing a type II error for $p=0.3$, $p=0.4$ and $p=0.5$

Try 10.8

One and two tailed test

A test of any statistical hypothesis, where the alternative is one sided, such as

$$H_0: \theta = \theta_0$$
$$H_1: \theta > \theta_0$$

or perhaps,

$$H_0: \theta = \theta_0,$$
$$H_1: \theta < \theta_1$$

is called a one tailed test.

On the other hand where the alternative is two sided, such as

$$H_0: \theta = \theta_0$$
$$H_1: \theta \neq \theta_0, \quad (\theta < \theta_0 \text{ or } \theta > \theta_0)$$

is called a two sided.

Example		
	10.1	→ One tail test.
	10.2	→ ~~Two~~ Two tail test.

Significance level:

If the probability of a value of the test statistic 'as bad as worse' as that obtained is p then we reject the null hypothesis if p is less than or equal to the significance level α.

" The significance level is the level of probability that we call unlikely. If your test gives a probability as unlikely as the significance level then you reject the null hypothesis. The usual significance level is 5% or 0.05 but other level such as 1% (0.01) or 10% (0.1) are often used.

Example:

4 - people playing cards

So ~~wining~~ P(wining each person) = $\frac{1}{4}$

Claim Sarah cheating.

$H_0 : p = \frac{1}{4}$ (innocent)

$H_1 : p > \frac{1}{4}$ (guility / not innocent)

If Sarah wins 5 out of 10 then proportion is $\frac{1}{2}$. (as bad or worse).

$X \sim B(10, \frac{1}{4})$

$$\therefore P(X \geq 5 \mid X \sim B(10, \tfrac{1}{4})) = 1 - P(X \leq 4 \mid X \sim B(10, \tfrac{1}{4}))$$
$$= 1 - 0.9219$$
$$= 0.0781$$

ie, 7.8% is reasonably large (more than 5%)

So there is no reason to suspect the validity of H_0 and Sarah remains innocent. There is insufficient evidence to reject the null hypothesis.

What value of x would provide sufficient evidence to reject the claim that $P = \frac{1}{4}$? ie, Sarah is guilty.

If we use a 5% significance level then we need the value of c such that

$$P(X \geq c) \leq 0.05 \quad \text{where } X \sim B(10, 0.25)$$

ie, ~~$1 - P(PX \leq c)$ ≤ 0.05~~

$$1 - P(X \leq c-1) \leq 0.05$$

$$P(X \leq c-1) \geq 0.95$$

$$P(X \leq 4) = 0.9219$$

$$P(X \leq 5) = 0.9803$$

ie, $c - 1 = 5$ or $c = 6$

we have $P(X \geq 6) = 0.0197 = 1.97\% \ (< 5\%)$

So if Sarah had won 6 or more games out of 10 we would had a significant result to reject H_0.

ie we reject H_0 at the 5% level of significance.

$X \geq 6 \rightarrow$ critical region

$6 \rightarrow$ cirtical value.