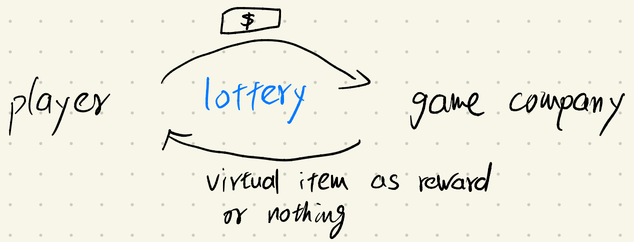
Appetizer: How unlucky is unlucky?



Rules

- $1 per shot. non-refundable
- prob $= \frac{1}{10}$ to win.
- player may shoot as many times, may stop any time

Outcome

- 30 shots total
- 1 win, 29 lose $\Rightarrow$ observed prob $= \frac{1}{30}$
- player sues the company

If you were the { judge / player / company }, what would you do?

$$P(1\text{ win } 29 \text{ lose} \mid p = \tfrac{1}{10}) = \binom{29}{1} \cdot \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^{29} \approx 0.137$$

Is this enough to prove the company guilty?
or the player is having bad luck?

P-value

It's used to test the null hypothesis H_0 (the initial hypothesis)

If $p < .05$, then we reject H_0

Otherwise we don't reject H_0

Discrete case

p-value = the prob. of seeing sth that's equally rare

+ the prob of seeing sth rarer or more extreme

+ the prob of seeing the observed

The prob calculated with H_0

eg.

H_0 = fair coin for 5 times   Is the coin fair?

result = 4 heads, 1 tail

$P(4H, 1T) = \frac{5}{32}$ $P(5H) = P(5T) = \frac{1}{32}$

$P(4T, 1H) = \frac{5}{32}$ $\therefore$ p-value $= \frac{5}{32} + \frac{5}{32} + \frac{1}{32} \times 2 = 0.375$

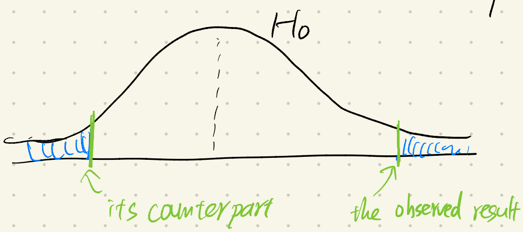
Fail to reject H_0

$\neq$ accept H_0

More data is needed.

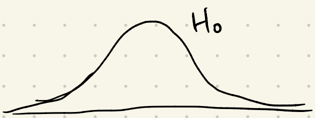
Continuous case

p-value = the blue area



If we reject Ho, then it means there is a better distribution to fit the data
Otherwise Ho is good enough

Approximate 5% of stats tests from the same distribution will be false positive



random sample
to calculate p-value → p = 0.68, 0.49, 0.92, 0.02

How tall is human?
I knew they're around 5-7 ft = Ho

Unfortunately, the alien went to a kindergarten,
and recorded avg = 2.6 ft, SD = 0.4 ft
⇒ p = 0.005 ⇒ reject Ho

⇒ it's a false positive

...Human are shorter than I thought

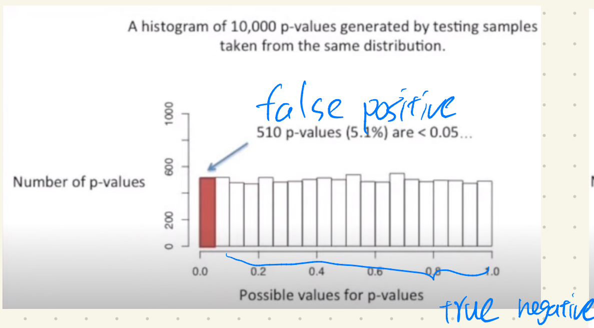
Confusion matrix

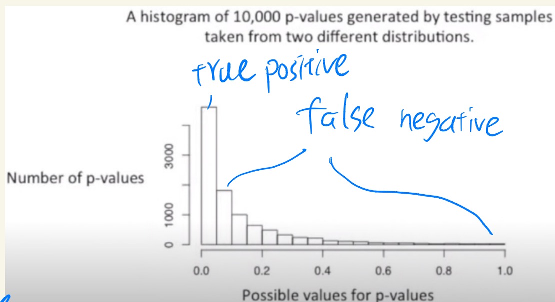
		Reject null hypothesis: No ($p > 0.05$)	Reject null hypothesis: Yes ($p < 0.05$)
Null hypothesis	True	True negative	False positive (type 1 error)
	False	False negative (type 2 error)	True positive

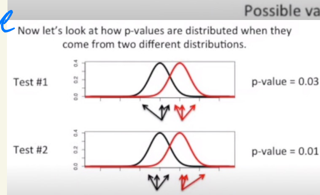
eg.
H_0 ~ healthy, Test ~ virus test
positive = reject H_0 = unhealthy
= test positive

$$\text{specificity} = \frac{TN}{TN + FP}$$

$$\text{sensitivity} = \text{recall} = \frac{TP}{TP + FN}$$







p-values are uniform if draw from the same distri.

p-values are skewed or close to zero if draw from diff. distri.

Likelihood

$L(\theta_0; x) = \text{Prob}(X = x \mid \theta = \theta_0) = f_X(x; \theta_0)$ ← PDF

$L(\theta) = L(\theta; x) = \prod_{i=1}^{n} f_i(x_i; \theta)$

x = the data, θ = the parameters, n = # of samples

score $\frac{d}{d\theta} \ell(\theta)$

loglikehood $\ell(\theta) = \log L(\theta)$

Score equation $\frac{d}{d\theta} \ell(\theta) = 0$ is to find the max $\ell(\theta)$

Observed information $I_0(\theta) = -\frac{d^2}{d\theta^2} \ell(\theta)$ quantifies the confidence in the MLE (max. likelihood estimation

Expected information $I(\theta) = E[I_0(\theta; X)]$

expected value

Variance $V(\hat{\theta}) \approx I(\theta)^{-1}$

If θ is multi-dim, $\vec{\theta}$, then score equ: $\frac{\partial}{\partial \theta_i} \ell(\vec{\theta}) = 0$

information: $I_{ij}(\vec{\theta}) = -\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell(\vec{\theta})$

Bayes Thm

posterior, likehood, prior

$$P(\theta \mid D) = \frac{P(D \mid \theta) P(\theta)}{P(D)}$$

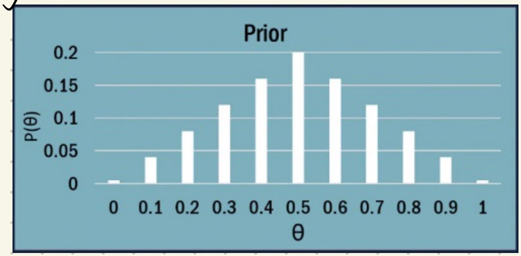
(normalizing const)

prior and posterior are distribution (pdf)

D = data

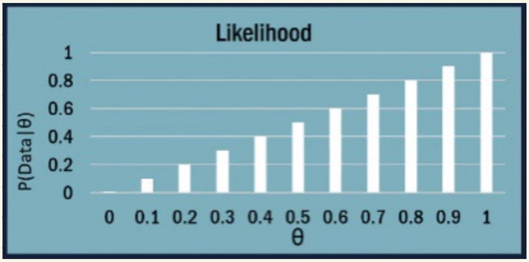
θ = model or parameter

eg. flip a coin



proportion of heads

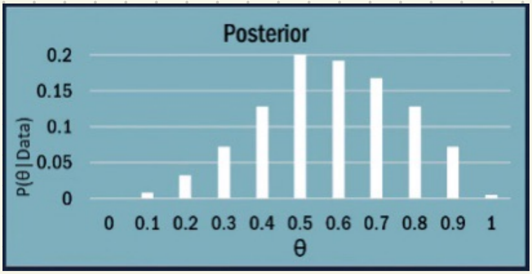
$P(\theta)$



let $n=1$, $head=1$

$P(D|\theta)$

*(not a pdf)



$P(\theta|D)$

If posterior has the same parametric form as prior
Then the prior is called conjugate prior

$$P(\theta|D) = \frac{P(D|\theta)\,P(\theta)}{\int P(D|\theta)\,P(\theta)\,d\theta}$$

Z test (Suppose σ is known)

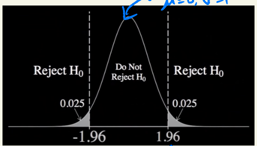
$H_0: \mu = \mu_0$,

known

Test stat : $Z = \dfrac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}$, where $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

If H_0 is true, then Z will be in normal distri

gaussian with $\mu=0, \sigma=1$



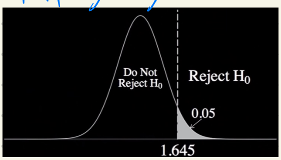
If $|Z| \geq 1.96$, then reject H_0 given $\alpha = 0.05$

Z value threshold use ppf to find

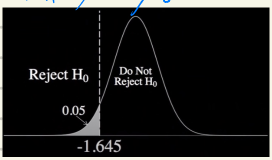
We don't calculate p-value

We convert α to the threshold

$H_1 : \mu > \mu_0$



$H_1 : \mu < \mu_0$



T test (Suppose σ is unknown)

$H_0 : \mu = \mu_0$

observed

Test stat : $t = \dfrac{\bar{x} - \mu_0}{SE}$, where $SE = \dfrac{s}{\sqrt{n}}$, s = SD of x

If H_0 is true, then t will be in t-distri with n-1 df

Chi-squared test

- goodness of fit
- test for indep

degree of freedom $df = (r-1)\cdot(c-1)$, $r = \#$ of rows, $c = \#$ of columns

$$\chi^2 = \sum_i \frac{(O_i - e_i)^2}{e_i}$$, O_i = observed, e_i = expected (calculated by H_0)

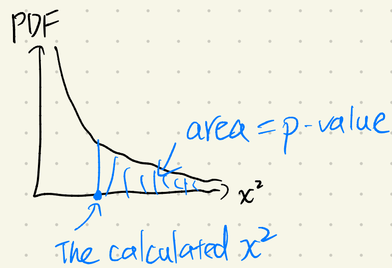
Reject H_0 if $\chi^2 >$ chi2.ppf(0.95, df) $= 3.84$

or if $p = 1 -$ chi2.cdf(χ^2, df) < 0.05

data → χ^2 → p-value ⎤
p-value threshold ⎦ compare

↓

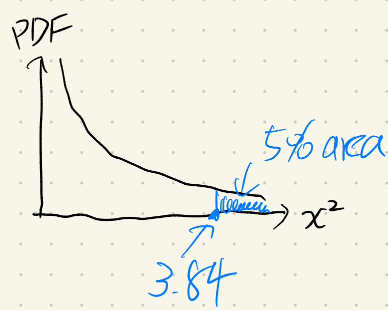
reject H_0 or not



p-value threshold → χ^2 threshold ⎤
data → χ^2 ⎦ compare

↓

reject H_0 or not



Inference for one variance

Let σ be SD of the population
s be SD of the samples

$H_0: \sigma = \sigma_0$

test stat: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$

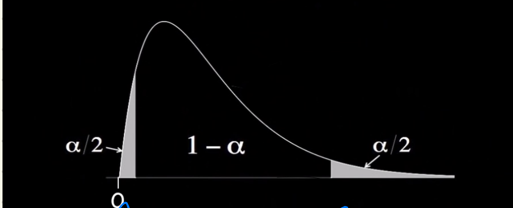
s^2 is unbiased estimator of σ^2, i.e. $E(s^2) = \sigma^2$

$\frac{(n-1)s^2}{\sigma^2}$ has χ^2-distri with $df = n-1$ (if the population is gaussian)

e.g. $\alpha = 0.05$

A $(1-\alpha)100\%$ confidence interval for σ^2 is given by:

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}\right)$$



$\chi^2_{1-\frac{\alpha}{2}}$

$\chi^2_{\frac{\alpha}{2}}$ use ppf to calculate

Calculate the confidence interval

$$P\left(\chi^2_{1-\frac{\alpha}{2}} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Rightarrow P\left(\underbrace{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}}_{\text{lower bound}} < \sigma^2 < \underbrace{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}}_{\text{upper bound}}\right) = 1-\alpha$$