MSE 213 – L7

$\boxed{1}$ RECIPE FOR z-TEST (GAUSSIAN TEST FOR THE MEAN)

1) Define $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{N_s}}$

2) Fixed probability? ↓ set P (e.g. 95%)

OR

Fixed value μ_0 / deviation z? ↓ determine z (e.g. 1.5)

3) Ask:

(a) Above which value μ_0...

(b) Below wich value μ_0...

(c) In which range $\bar{x} \pm |\mu_0 - \bar{x}|$

... is the true mean that represents my distribution, with probability P?

4) Answer:

(a) $\Phi(+z) = P$ (b) $\Phi(-z) = P$

(c) $2\Phi(|z|) - 1 = \Phi(|z|) - \Phi(-|z|) = P$

→ Find z from table.

→ Compute μ_0.

3) Ask:

With which probability is the true mean ~~of~~ that represents my population...

(a) above μ_0?

(b) below μ_0?

(c) in the range $\bar{x} \pm |\mu_0 - \bar{x}|$?

4) Answer:

(a) $P = \Phi(z)$ (b) $P = \Phi(-z) = 1 - \Phi(z)$

(c) $P = 2\Phi(|z|) - 1 = \Phi(|z|) - \Phi(-|z|)$

MSE 213 - L7

$\boxed{2}$ Two-sample Gauss / z-test.

- We are interested in the <u>difference</u> of two independent distributions X and Y. · We assume σ_X and σ_Y are known.

Distribution of the difference: Mean: $\bar{X} - \bar{Y}$

Variance: $V(X) + V(Y)$

(as $V(-Y) = (-1)^2 V(Y)$)

and is still Gaussian / Normal.

So we now use $z = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{\sigma_X^2}{N_{SX}} + \dfrac{\sigma_Y^2}{N_{SY}}}}$ and proceed as before.

$\boxed{3}$ One-sample Student t-test

We do not know σ, but estimate it from the unbiased estimator for σ, $s = \sqrt{\dfrac{1}{N_S - 1} \sum_{i=1}^{N_S} (x_i - \bar{x})^2}$.

We then get the "t-value" $t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{N_S}}$ (like z, but $\sigma \to s$).

We now have an estimator (s) in our equation, so the "degrees of freedom" (df) for our system are reduced by 1. $\boxed{df = N_S - 1}$

t now follows a "T-distribution" which is equal to a Gaussian ~~but~~ for $df \to \infty$ but broader (implying more uncertainty) for small df.

We need to look up its CDF in a "t table" now, which depends on df, but otherwise the recipe is the same as in $\boxed{1}$.

$\boxed{4}$ Two-sample t-test. Assuming both samples have the same variance, ~~[scribbled-out word]~~ ($\sigma_x = \sigma_y$) and are independent.

The "pooled" sample variance estimator is

$$S = \sqrt{\frac{1}{(N_{sx}-1)+(N_{sy}-1)}\left(\sum_{1}^{N_{sx}}(x_i-\bar{x})^2+\sum_{1}^{N_{sy}}(y_i-\bar{y})^2\right)}$$

$\bar{x}$ and $\bar{y}$ can be different so we substract separately!

Now $$t = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S^2}{N_{sx}}+\frac{S^2}{N_{sy}}}}$$ and $df = (N_{sx}-1)+(N_{sy}-1)$

and proceed as before.

E.g. we measured two independent groups of people who tried out different medicine.