$(X_1, \ldots, X_n)$ I.I.D. $\sim F(\mu_1, \sigma_1^2)$

$(Y_1 \ldots Y_m)$ I.I.D. $\sim F(\mu_2, \sigma_2^2)$

$F(x, \theta)$

$\theta = \mu, \sigma$

$\hat{\theta} = T(x_1 \ldots x_n)$

$\bar{X} = \frac{\sum x_i}{n} \rightarrow \mu$

$E[\bar{X}]$

θ – Just general notation for any parameter of a distribution (could be expectation, variance, proportion, etc)

Statisticians use it, when they need to talk about formulas/theory in general, without specifically emphasizing the parameter.

CLT:

$$\bar{X} = \frac{\sum X_i}{n}$$

$n > 30$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

Important consequence of CLT:

$X_1 \ldots X_n \quad K \qquad \hat{p}_1 = \frac{K}{n}$

$Y_1 \ldots Y_m \quad r \qquad \hat{p}_2 = \frac{r}{m}$

$$\hat{P} \sim N\left(P, \frac{P\cdot(1-P)}{n}\right), \quad n > 30$$

$$\hat{p}_2 \sim N\left(p_2, \frac{p_2(1-p_2)}{m}\right)$$

$$1-\alpha = P(L < \overset{\theta}{p_1 - p_2} < U)$$

$$\hat{\theta} = \hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}\right)$$

$$Z = \frac{X-\mu}{\sigma}$$

$$E[X+Y] = E[X] + E(Y)$$

$$Var(X-Y) = Var(X) + Var(Y)$$

$$Z$$

$$1-\alpha = P\left(\frac{\hat{\theta} - U}{-\|-} < \frac{\hat{\theta} - \theta}{\sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}}} < \frac{\hat{\theta} - L}{-\|-}\right)$$

$n = 94_m, \ {}^{\nearrow 30}$ 50 assistance p_m $\hat{p}_m = \frac{50}{94} = 0.53$

$K = 68\ F, \ {}_{\searrow 30}$ 40 assistance p_F $\hat{p}_F = \frac{40}{68} = 0.59$

$$1-\alpha = P(L < p_m - p_F < U)$$

$$L = \hat{p}_m - \hat{p}_F - z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_m(1-\hat{p}_m)}{n} + \frac{\hat{p}_F(1-\hat{p}_F)}{K}} = -0.06 - 2.576 \cdot 0.08$$

$$U = \hat{p}_m - \hat{p}_F + \text{—}//\text{—} \qquad U = -0.06 + 2.576 \cdot 0.08$$

Solution from the seminar with another group

$$L = \hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}$$

$$U = \hat{p}_1 - \hat{p}_2 + \text{———} \,//\, \text{———}$$

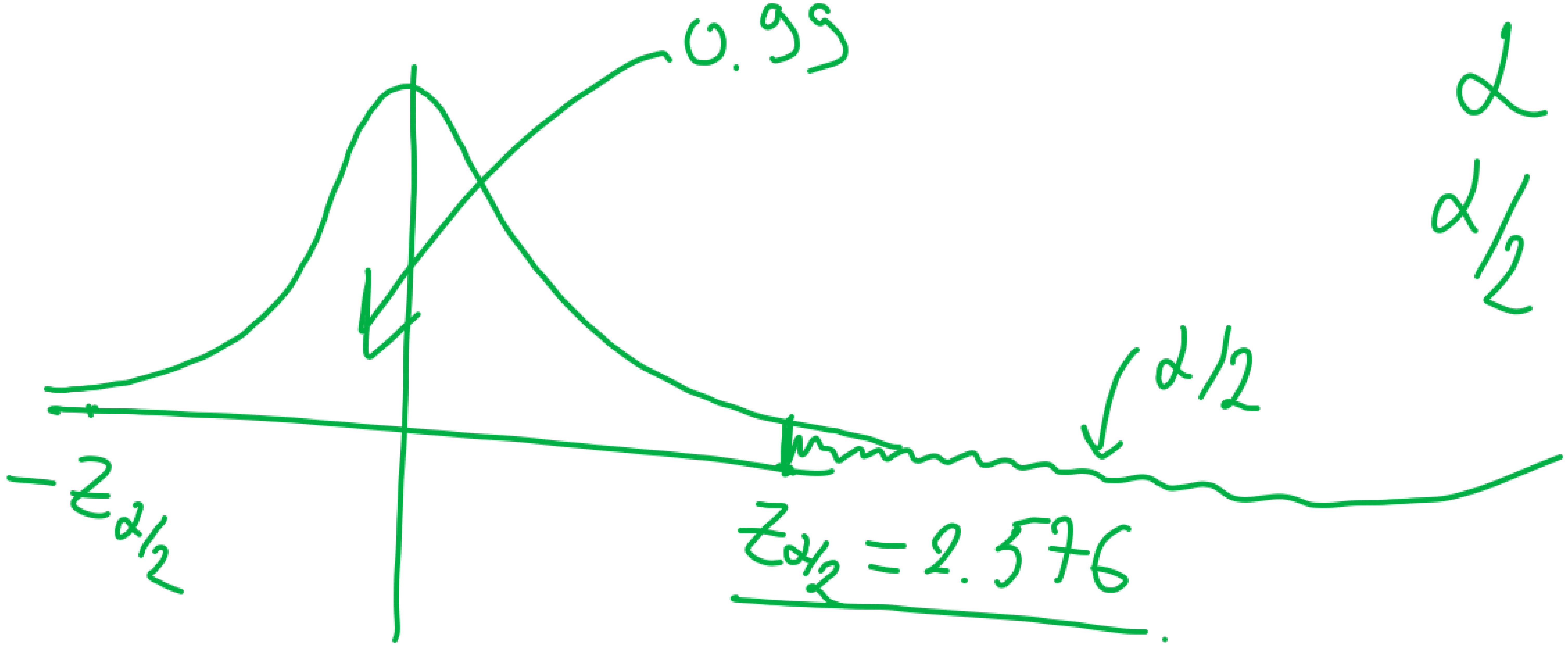
Ex: 94 ; 50
n ; k

68 ; 40
m ; r

$p_m - p_F$

$$\hat{p}_m = \frac{50}{94} = 0.53$$

$$\hat{p}_F = \frac{40}{68} = 0.59$$



$\alpha = 0.01$

$\alpha/2 = 0.005$

$$P_1 - P_2 \in \left(-0.06 - 2.576\sqrt{\frac{\hat{P}_m(1-\hat{P}_m)}{94} + \frac{\hat{P}_F(1-\hat{P}_F)}{68}}\ ,\ -0.06 + \text{—//—} \right)$$

$$P_1 - P_2\ (-0.06 - 0.02,\ -0.06 + 0.02)$$

Confidence Intervals using Student's t-distribution

1. For the cases when population variance is unknown
2. Requirement: Number of observations in a sample should be greater than 30,
   or sample should be taken from the normally distributed population

3. Meet new characters: Chi-squared distribution, Student's t-distribution

μ — ?

σ^2 - known

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$E[S^2] = \sigma^2$$

$$\frac{S^2(n-1)}{\sigma^2} \sim \chi^2_{(n-1\,df)}$$

$\sim N$, $\sim \chi^2$

$$\chi^2(k) = \sum_{i=1}^{k} Z_i^2$$

, k - degrees of freedom

$$T_{(k)} = \frac{Z}{\sqrt{\frac{\chi^2(k)}{k}}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{S^2(n-1)}{\sigma^2(n-1)}}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1\,df)}$$

$$1-\alpha = P(L < \mu < U)$$

$$1-\alpha = P\left(\underbrace{\frac{\bar{X} - U}{\frac{S}{\sqrt{n}}}}_{-t_{\alpha/2}} < \underbrace{\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}}_{t_{(n-1\ d.f)}} < \underbrace{\frac{\bar{X} - L}{\frac{S}{\sqrt{n}}}}_{t_{\alpha/2}} \right)$$

$$\boxed{\begin{aligned} U &= \bar{X} + t_{\alpha/2} \cdot \frac{S}{\sqrt{n}} \\ L &= \bar{X} - t_{\alpha/2} \cdot \frac{S}{\sqrt{n}} \end{aligned}}$$

① $X_1 \dots X_5 \sim N(\mu, \sigma^2)$ $\quad m=5$

$\bar{X} = 100$

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{m}\right)$ $\quad$ 90% CI

$X_1^{(2)} \dots X_{10}^{(2)}$; $S^2 = 9$

$$T_{(n-1\ df)} = \frac{Z}{\sqrt{\frac{\chi^2(k)}{k}}} = \frac{\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{m}}}}{\sqrt{\frac{S^2(n-1)}{\sigma^2\cdot(n-1)}}} = \frac{\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{5}}}}{\sqrt{\frac{9}{\sigma^2}\cdot\frac{9}{9}}} = \frac{\bar{X}-\mu}{\sqrt{9/5}}$$

$$1-\alpha = P\left(\frac{\bar{X}-U}{\frac{S}{\sqrt{n}}} < t_{(n-1)} < \frac{\bar{X}-L}{\frac{S}{\sqrt{n}}} \right)$$

$$U = \bar{X} + t_{\alpha/2} \cdot \frac{S}{\sqrt{n}} = 100 + 1.833 \cdot \frac{3}{\sqrt{5}}$$

$$L = 100 - 1.833 \frac{3}{\sqrt{5}}$$