本课程的主要内容:

§9.4 两个正态总体的假设检验;

§9.5 总体分布的假设检验;

§9.4 两个正态总体的假设检验

一、两个正态总体均值差的假设检验（独立样本）

(一)已知方差σ_1^2,σ_2^2，检验$H_0:\mu_1=\mu_2$（大.小样本）

选用统计量: $$U=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

由: $P(|U|>\lambda)=\alpha$ 得: $\Phi(\lambda)=1-\frac{\alpha}{2}$ $\lambda=z_{\frac{\alpha}{2}}$

得拒绝域: $|U|>\lambda$

若在大样本下，方差 σ_1^2, σ_2^2，未知

可选择统计量:

$$U = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$$

由: $P(|U| > \lambda) = \alpha$ 得: $\Phi(\lambda) = 1 - \frac{\alpha}{2}$ $\lambda = z_{\frac{\alpha}{2}}$

得拒绝域: $|U| > \lambda$

例:某企业对男女职员的平均小时工资进行了调查,独立抽取了具有同类工作经验的男女职员的两个随机样本,并统计下两个样本的均值、方差等数据:

男职员: $n_1 = 44, \bar{x}_1 = 75元, s_1^2 = 64$

女职员: $n_2 = 32, \bar{x}_2 = 70元, s_2^2 = 42.25$

在显著水平为0.05的条件下,能否定为男性职员与女性职员的平均小时工资存在显著差异?

解: $H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 \neq 0,$

选择统计量: $$U = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$$

由 $P\left(|U|>\lambda\right)=\alpha$ 得临界值: $\lambda=1.96$

拒绝域为: $|U|>1.96$　　由样本计算:

$$|U|=\left|\frac{75-70}{\sqrt{\frac{64}{44}+\frac{42.25}{32}}}\right|=3.002>1.96$$

结论:拒绝H_0,能够以为该企业男女职员的平均小时工资之间存在明显差别.

（二）共同总体原则差的小样本检验

未知方差σ_1^2,σ_2^2，但$\sigma_1^2=\sigma_2^2$，检验$H_0:\mu_1=\mu_2$

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$$

其中$S_w=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

(混合样本原则差)

查t分布表，得临界值$\lambda=t_\alpha(n_1+n_2-2)$

得拒绝域: $|U|>\lambda$

例:甲乙两台机床同步加工某种同类型零件,已知两台机床加工的零件直径(单位:cm)分别服从正态分布$N_{\circ}\left(\mu_1,\sigma^2\right),N\left(\mu_1,\sigma^2\right)$ 为比较两台
度有无明显差别,分别独立抽取了甲机床加工的8个零件和乙加工的7个零件,经过测量得到数据:

甲:20.5 19.8 19.7 20.4 20.1 20.0 19.0 19.9

乙:20.7 19.8 19.5 20.8 20.4 19.6 20.2

在明显水平0.05下,样本数据是否提供证据支持"两台机床加工的零件直径不一致"的看法?

解: $H_0: \mu_1-\mu_2=0, H_1: \mu_1-\mu_2\neq 0,$

选择统计量: $T=\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$

由 $P(|T|>\lambda)=\alpha$ 得临界值: $t_{0.05}(13)=2.106$

拒绝域: $|T|>2.106$ 由样本计算:

$\bar{x}_1=19.925, \quad \bar{x}_2=20.143. \quad s_1^2=0.2164, \quad s_2^2=0.2729$

$$S_w = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = 0.2425$$

$$|T| = \left| \frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \right| = 0.855 < 2.160$$

结论:没有理由拒绝H_0, 以为:甲乙机床加工的零件平均直径不存在明显差别.

二两个正态总体方差比的假设检验(独立样本)

未知μ_1,μ_2,检验假设H_0: $\sigma_1^2=\sigma_2^2$

（总体方差的齐性检验）

选择统计量 $F=\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$

由: $P(F<\lambda_1,F>\lambda_2)=\alpha$ 查F分布表可得临界值

$\lambda_1=F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$ $\lambda_2=F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$

得拒绝域: $F<\lambda_1,F>\lambda_2$

例:在10个相同的地块上对甲、乙两种玉米进行品比试验,得数据(单位:公斤)

甲：951　966　1008　1082　983

乙: 730　864　742　774　990

假定农作物产量服从正态分布,在明显水平0.05下,检验两个总体的均值是否相等,以及方差是否相等。

解:第一步检验方差的齐性, 即 H_0: $\sigma_1^2=\sigma_2^2, H_1$: $\sigma_1^2 \neq \sigma_2^2$,

选择统计量 $F=\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}=\dfrac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1)$

由 $P(F>\lambda_2)=\dfrac{\alpha}{2}, P(F<\lambda_1)=\dfrac{\alpha}{2}$ 得临界值

$$\lambda_2 = F_{0.025}\left(5-1,5-1\right) = 9.6$$

$$\lambda_1 = F_{1-0.025}\left(5-1,5-1\right) = 1/F_{0.025}\left(5-1,5-1\right) = 1/9.6 = 0.1$$

由样本 $T = \dfrac{S_1^2}{S_2^2} = \dfrac{2653}{11784} = 0.23$ 结论:不能拒绝H_0.

第二步检验均值是否存在明显差别

$$H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 \neq 0,$$

选择统计量: $T = \dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$

由 $P(|T| > \lambda) = \alpha$ 临界值: $t_{0.05}(8) = 2.306$

拒绝域: $|T| > 2.306$ 由样本计算:

$$\bar{x}_1 = 998, \quad \bar{x}_2 = 820. \quad s_1^2 = 2653.5, \quad s_2^2 = 11784$$

$$S_w = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 7218.75$$

$$|T| = \left| \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| = 3.313 > 2.306$$

单侧检验P184例3

结论:拒绝H_0, 以为两种玉米地存在明显差别.

三.成对数据均值的假设检验（不独立样本）

例:某饮料企业开发研制出一新产品,为比较消费者对新老产品口感的满意程度,该企业随机抽选一组消费者8人,每个消费者先品尝一种饮料,然后再品尝另一种饮料,两种饮料的品尝顺序是随机的,而后每个消费者要对两种饮料分别进行评分0~10分,成果如下:

旧:	5	4	7	3	5	8	5	6
新:	6	6	7	4	3	9	7	6

在明显水平0.05下检验消费者对两种饮料的口感是否存在明显差别?