第 2 章　高维列联表资料的常规统计分析

当观测结果是定性资料时，人们习惯将资料整理成列联表形式，如 2 ×2 列联表资料、$R \times C$ 列联表资料和高维列联表资料等。所谓高维列联表，也就是表中涉及的定性变量的个数 $k \geqslant 3$。对于高维列联表资料，根据结果变量的性质可将高维列联表分为三类：一是结果变量为二值变量的高维列联表；二是结果变量为多值有序变量的高维列联表；三是结果变量为多值名义变量的高维列联表。本章将详细介绍高维列联表资料的常规统计分析方法及其 SAS 软件实现。

2.1　问题与数据

【例 2-1】 为调查眼科门诊患者的用药依从性及其影响因素，并提出相应对策，提高眼科门诊药房的药学服务水平，在 5 家医院的眼科门诊患者中，随机发放 610 份调查问卷，进行汇总，数据见表 2-1，试分析不同告知者之间患者的依从率有无差别。

表 2-1　不同告知者、不同告知程度间患者的依从情况

告知者	告知程度	调查人数	依从人数	不依从人数	依从率(%)
医生	详细告知	264	240	24	90.91
	告知	262	202	60	77.10
	没告知	47	31	16	65.96
药师	详细告知	207	188	19	90.82
	告知	298	238	60	79.86
	没告知	68	47	21	69.12

【例 2-2】 有 7 项关于阿司匹林预防心肌梗死后死亡的随机对照实验研究，结果见表 2-2，试对该资料进行 meta 分析。

表 2-2　阿司匹林预防心肌梗死后死亡的 7 项随机对照实验研究结果

研究编号	观察人数(死亡人数)		研究编号	观察人数(死亡人数)	
	阿司匹林组	安慰剂组		阿司匹林组	安慰剂组
S1	615(49)	624(67)	S5	810(85)	406(52)
S2	758(44)	771(64)	S6	2267(246)	2257(219)
S3	832(102)	850(126)	S7	8578(1570)	8600(1720)
S4	317(32)	309(38)			

【例 2-3】 调查某中医院一日内医生开出的针对甲、乙两种疾病的处方情况，结果见表 2-3，试分析不同疾病、不同性别的患者所用药物种类频数构成有无差别。

【例 2-4】 某研究为评价某近视人群的视力水平，描述此近视人群特征，对近视度数进行调查，数据见表 2-4，试分析此近视人群性别间、不同年龄段的视力水平有无差别。

表 2-3　不同疾病、不同性别与药物种类条件下受试者的频率的关系

疾病类型	患者性别	例数		
	药物种类：	中药	西药	复方药
甲	男	52	7	2
	女	34	8	1
乙	男	23	19	4
	女	18	11	3

表 2-4　某近视人群近视度数调查结果

性别	年龄段	例数		
	近视度数：	<3.00 D	3.00～6.00 D	>6.00 D
男	30～39	10	26	10
	40～49	8	22	5
	≥50	5	13	2
女	30～39	20	6	18
	40～49	22	10	15
	≥50	36	1	1

2.2　对数据结构的分析

对于例 2-1 的资料，涉及 3 个定性变量，即“告知者”、“告知程度”和“是否依从”，结果变量“是否依从”是一个二值结果变量，数据以列联表的形式呈现，故该资料属于结果变量为二值变量的高维列联表资料。

例 2-2 中各项原始研究所对应的试验设计类型均为单因素两水平设计（或成组设计），而且属于结果变量为二值变量的前瞻性研究，因此，在流行病学上又将此类研究称为队列研究，将所得到的资料称为队列研究设计四格表资料。表 2-2 中的每一行都是一个 2×2 的四格表资料，故该表实际可整理为结果变量为二值变量的高维列联表的形式。

例 2-3 中含有三个定性变量，分别为“疾病类型”、“患者性别”、“药物种类”，结果变量为“药物种类”（多值名义变量），因此该表称作结果变量为多值名义变量的三维列联表。

例 2-4 中含有三个变量，分别为“性别”、“年龄段”、“近视度数”，结果变量为“近视度数”（多值有序变量），因此该表称作结果变量为多值有序变量的高维列联表。

2.3　统计分析目的与分析方法选择

对例 2-1 而言，通常可选 CMH 校正的χ^2 检验（也称加权χ^2 检验）、对数线性模型和多重 logistic 回归模型分析。CMHχ^2 检验可以控制一个或多个对定性结果变量影响相对次要的因素（或变量）而重点考察一个研究者最关心的影响因素（即变量）对定性结果变量的影响情况。若希望用较简便的统计分析方法评价不同告知者的患者依从效果，并尽量消除“告知程度”对结果的影响，可使用 CMH 校正的χ^2 检验。

例 2-2 中各项研究所对应的试验设计类型均为队列研究设计，故此表资料本质上为 7 个队列研究设计的四格表（或 2×2 表）资料，可对该资料进行 meta 分析（当然，也可进行 CMH 校正的χ^2 检验，只是后者没有像 meta 分析那样给出更多详细的分析而已）。

对于例 2-3 中的结果变量为多值名义变量的高维列联表资料，分析目的是考察不同疾病、不同性别的患者所用药物种类频数构成有无差别，分析时通常可选用 CMH 校正的χ^2 检验、扩展的多重 logistic 回归分析或对数线性模型，这里采用 CMH 校正的χ^2 检验进行分析。

对于例 2-4 中的结果变量为多值有序变量的高维列联表资料，分析时通常可选用 CMH 校正的秩和检验分析或有序变量的多重累计 logistic 回归模型分析。

2.4　用 SAS 软件对实例进行分析

2.4.1　用 SAS 软件对例 2-1 的资料进行加权χ^2 检验

表 2-1 看似是该研究问题的“标准型”，但实质上还是欠完善的。这是因为告知程度在研究中是要作为一个“分层因素”的，研究的目的就是要尽量排除该分层因素对结果产生的影响，故将该资料整理成适合用加权χ^2 检验的统计表，见表 2-5。

表 2-5　不同告知者、不同告知程度间患者的依从情况

告知程度	告知者	例数	
		依从	不依从
详细告知	医生	240	24
	药师	188	19
告知	医生	202	60
	药师	238	60
没告知	医生	31	16
	药师	47	21

对例 2-1 进行加权χ^2 检验，SAS 程序(程序名为 TJFX2_1A. SAS)如下：

```
data tjfx2_1A;
do k=1 to 3;
input a b c d;
n=a+b+c+d; e=a+b; f=c+d;
g=a+c; h=b+d;
d1=(a* d-b* c)/n;i=c+d;
d2=e* f/n;  pd=e* f* g* h; v1=pd/n* * 3;k2=pd/(n-1)/n* * 2;
j=a* f;k=c* e;
l=j/n;m=k/n;output; end;
cards;
240  24  188  19
202  60  238  60
31   16   47   21
;
run;
ods html;
proc means noprint;
var d1 d2 v1 k2 l m;
output out=aaa sum=d1 d2 v1  k2 l m; run;
data a; set aaa;
v2=d2* * 2; md=d1/d2; v=v1/v2; k1=d1* * 2;
```

```
wchisq=ROUND(md* * 2/v, 0.001); wp=2 - PROBCHI(wchisq, 1);
rr=round(1/m, 0.001);  mhchisq=ROUND(k1/k2, 0.001);
mhp=2 - PROBCHI(mhchisq, 1);
file print;
put #3  @ 5 'W - chisq=' wchisq  @ 35 'W - p=' wp
#6  @ 5  'RR=' rr @ 15 'MH - chisq=' mhchisq @ 35 'MH - p=' mhp; run;
ods html close;
```

【程序说明】 该程序的主体是通过对公式进行编程来计算加权χ^2和MHχ^2统计量及RR值。

【输出结果及解释】

W-chisq=0.548　　W-p=0.4591360465

RR=0.98　　MH-chisq=0.547　　MH-p=0.45954609

统计结论：W－chisq为加权χ^2值，W－p为对应的P值，即加权$\chi^2=0.548$，$P=0.4591$，说明消除了告知程度对结果的影响后，医生与药师对应的患者的依从率之间的差别无统计学意义；相对危险度为RR=0.98，说明医生对应的患者的依从率是药师对应的患者的依从率的0.98，即医生对应的患者的依从率略低于药师对应的患者的依从率。在总体上，RR是否等于1？回答此问题的检验方法以前被称为MHχ^2检验，后来被称为CMH校正的χ^2检验，在本例中，MH$\chi^2=0.547$，$P=0.4595$，说明在总体上，RR几乎等于1。

专业结论：消除告知程度的影响后，告知者之间患者依从率的差别没有统计学意义。医生组的患者依从率为$P_e=a/(a+b)$，药师组的患者依从率为$P_0=c/(c+d)$，则$RR=P_e/P_0=0.98$，此值与1之间的差别无统计学意义，说明告知者分别是医生与药师对应的患者的依从率基本相同。

2.4.2　用SAS软件对例2-1资料进行CMH校正的χ^2检验

对例2-1进行CMH校正的χ^2检验，SAS程序(程序名为TJFX2_1B.SAS)如下：

```
data tjfx2_1B;
    do cd=1 to 3;
        do yy=1 to 2;
            do yc=1 to 2;
                input f @@ ;
                output;
            end;
        end;
    end;
cards;
240  24  188  19
202  60  238  60
31   16  47   21
;
run;
ods html;
proc freq;
    weight f;
     tables cd* yy* yc / cmh score =
rank;
run;
ods html close;
```

【程序说明】 cd、yy 和 yc 分别代表“告知程度”、“告知者”和“依从与否”；cmh 要求进行 CMH 校正的χ^2 检验；score = rank 要求对定性的结果变量的每档按其所在档次中的平均秩赋值，而不是按表得分(该程序中，yc = 1 to 2 就是给依从赋值为 1 分，给不依从赋值为 2 分)赋值。

【输出结果及解释】

Cochran- Mantel- Haenszel 统计量(基于秩得分)				
统计量	对立假设	自由度	值	概率
1	非零相关	1	0.5152	0.4729
2	行均值得分差值	1	0.5104	0.4750
3	一般关联	1	0.5466	0.4597

以上结果中，应看最后一行，H_0：行变量(告知者)与列变量(依从与否)之间互相独立，H_1：行变量(告知者)与列变量(依从与否)之间存在一般关联。$\chi^2_{CMH} = 0.547$，$P = 0.4597$(注意，此结果实际上就是前面第 1 段 SAS 程序输出的“MH - chisq = 0.547，MH - p = 0.4595”，它实际上是检验总体 RR 是否为 1，而不是检验行变量与列变量是否独立，SAS 软件试图用此一个检验取代“H_0：行变量与列变量相互独立”和“H_0：RR = 1”两个检验，虽然有点欠妥，但这两个检验的结果十分接近)。

普通相对风险的估计值(行 1/行 2)				
研究类型	方法	值	95% 置信限	
案例对照	Mantel- Haenszel	0.8887	0.6499	1.2152
(优比)	Logit	0.8886	0.6500	1.2148
Cohort	Mantel- Haenszel	0.9804	0.9300	1.0335
(第 1 列风险)	Logit	0.9888	0.9431	1.0367
Cohort	Mantel- Haenszel	1.0980	0.8575	1.4059
(第 2 列风险)	Logit	1.1014	0.8611	1.4089

以上结果中，按告知程度进行了分层，每层都是一个四格表资料，因为先由告知者发出“指令”，后由患者作出反应(即是否依从)，故可视为“队列(Cohort)研究设计”。与其对应的有 4 行，前两行(标注“第 1 列风险”)的含义是：以第 1 列为研究者关心的结果(即“依从”)来计算两行(医生与药师)上的相对危险度 RR 的估计值(即医生对应的患者的依从率除以药师对应的患者的依从率)，此时，又有 Mantel- Haenszel 算法与 Logit 算法两种算法计算的相对危险度 RR 的估计值及其 95% 置信区间。通常，取 Mantel- Haenszel 算法给出的结果，即 RR = 0.98，其 95% 置信区间为(0.9300，1.0335)。

对优比的齐性的 Breslow- Day 检验	
卡方	0.2135
自由度	2
Pr > 卡方	0.8988

以上结果中，显示了按告知程度分层后，各层算出的 RR 值是否相等的检验结果，即采用 Breslow- Day 检验，得到$\chi^2_{B-D} = 0.2135$，$df = 2$，$P = 0.8988$，说明三层中的 RR 接近相等，即相对危险度(标注的是优势比，此表述适用于病例对照研究设计的场合)满足齐性要求，这是

可否进行前面的 CMH 校正的χ^2 检验的前提条件。严格地说，只有在 Breslow – Day 检验得出 $P>0.05$ 的条件下，前面的 CMH 校正的χ^2 检验的结果才是可用的。

统计结论：由上述结果可以看出在控制了 b 因素(即“告知程度”)后，反映一般关联性高低的卡方值为 0.5466，$P=0.4597>0.05$，表明告知者之间患者依从率的差别没有统计学意义。

专业结论：可以认为不同告知者所对应的患者依从率接近相等。

值得注意的是：在前面的两小节中，一共用两段 SAS 程序实现了计算。2.4.2 小节的 SAS 程序中使用了 FREQ 过程和 CMH 选项，给出的结果与 2.4.1 小节 SAS 程序的结果略有不同，即“W – chisq = 0.548，W – p = 0.4591”在 2.4.2 小节的 SAS 程序的输出结果中是没有的，它就是所谓的“加权$\chi^2_{加}$ 检验及其结果”，目的是检验“ H_0：行变量与列变量相互独立”这个零假设。

2.4.3　用 SAS 软件对例 2-2 的资料进行 meta 分析

对例 2-2 进行 meta 分析，SAS 程序(程序名为 TJFX2_2.SAS)如下：

```
options ls=74 ps=max center nodate;
data tjfx2_2;
input study_id ni1 n_ai ni2 n_ci @@ ;
n_bi=ni1-n_ai;n_di=ni2-n_ci;
n_ai+0.25;n_bi+0.25;n_ci+0.25;n_di+0.25;
rr=(n_ai/(n_ai+n_bi))/(n_ci/(n_ci+n_di));
yi=log(rr);
wi=(1/n_ai-1/(n_ai+n_bi)+1/n_ci-1/(n_ci+n_di))**(-1);
ni=n_ai+n_bi+n_ci+n_di;
number_of_study=7;
cards;
1  615   49   624   67
2  758   44   771   64
3  832  102   850  126
4  317   32   309   38
5  810   85   406   52
6  2267  246  2257  219
7  8578 1570  8600 1720
;
run;
data b;set tjfx2_2;
wi2=wi**2;wiyi=wi*yi;wiyi2=wi*yi**2;
sum_of_wi_temp+wi;sum_of_wi2_temp+wi2;
sum_of_wiyi_temp+wiyi;
sum_of_wiyi2_temp+wiyi2;
id=_n_;
if _n_=number_of_study then do;
sum_of_wi=sum_of_wi_temp;
sum_of_wi2=sum_of_wi2_temp;
sum_of_wiyi=sum_of_wiyi_temp;
sum_of_wiyi2=sum_of_wiyi2_temp;
yw_bar=sum_of_wiyi/sum_of_wi;
square_of_s_y_bar=1/sum_of_wi;
q=sum_of_wiyi2-yw_bar**2*sum_of_wi;
df=number_of_study-1;
p=1-probchi(q,df);
i2=(q-(number_of_study-1))/q*100;
```

```
if i2 <0 then i2 =0;
frrc =exp(yw_bar);
frrlow =exp(yw_bar -1.96* square_of_s_y_bar* * 0.5);
frrup =exp(yw_bar +1.96* square_of_s_y_bar* * 0.5);w_bar =sum_of_wi/number_of_
study;
square_of_s_w =(sum_of_wi2 -number_of_study* w_bar* * 2)/(number_of_study
-1);
u =(number_of_study -1)* (w_bar -square_of_s_w/(number_of_study* w_bar));
square_of_tao_hat =(q -(number_of_study -1))/u;
if square_of_tao_hat <0 then square_of_tao_hat =0;end;
run;
proc sort;by descending id;
data c;set b;
square_of_tao_hat_for_all +square_of_tao_hat;
wi_bar =1/(1/wi +square_of_tao_hat_for_all);
run;
proc sort;by id;
data d;set c;
sum_of_wi_bar_temp +wi_bar;
sum_of_wi_baryi_temp +wi_bar* yi;
if _n_ =number_of_study then do;
sum_of_wi_bar =sum_of_wi_bar_temp;
sum_of_wi_baryi =sum_of_wi_baryi_temp;
v_hat_y_bar_hat =1/sum_of_wi_bar;
y_bar_hat =sum_of_wi_baryi/sum_of_wi_bar;
rrc =exp(y_bar_hat);
rrlow =exp(y_bar_hat -1.96* v_hat_y_bar_hat* * 0.5);
rrup =exp(y_bar_hat +1.96* v_hat_y_bar_hat* * 0.5);end;
run;
proc sort;by descending id;
data e;set d;where id =number_of_study;run;
ods html;
proc print;
var number_of_study q df p i2 frrc frrlow frrup rrc rrlow rrup;
run;
ods html close;
```

【程序说明】 该程序中分别用 ni1、ni2 表示各研究中试验组和对照组的样本含量；用 Ni 表示各研究的样本含量；分别用 n_ai、n_bi、n_ci、n_di 表示各研究对应四格表 4 个格子上的频数；用 yi 表示相对危险度 RR 的对数值；用 wi 表示采用固定效应模型时的权重；用 wi2 表示 wi 的平方；用 wiyi 表示 wi 与 yi 的乘积；用 wiyi2 表示 wi 与 yi 的平方的乘积；用 number_of_study 表示纳入研究的数量；用 Sum_of_wi_temp、Sum_of_wi2_temp、Sum_of_widi_temp 和 Sum_of_widi2_temp 分别表示 wi、wi2、widi 和 widi2 累积的中间结果，用 Sum_of_wi、Sum_of_wi2、Sum_of_widi 和 Sum_of_widi2 分别表示 wi、wi2、widi 和 widi2 的累积的最终结果；用 yw_bar 和 y_bar_hat 分别表示采用固定效应模型和随机效应模型时的合并效应；用 w_bar、Square_of_s_w、Square_of_tao_hat 和 wi_bar 分别表示 $\bar{w}$ 、s_w^2 、$\hat{\tau}^2$ 和 $\bar{w}_i$ ；用 DF 表示自由度；用 Q 和 P 分别表示异质性检验的 χ^2 值和 P 值；I2 为异质性检验的另一个统计量；用 Sum_of_wi_bar_temp 和 Sum_of_wi_baryi_temp 分别表示 wi_bar 和 wi_baryi 累积的中间结果，用 Sum_of_wi_bar 和 Sum_of_wi_baryi 分别表示 wi_bar 和 wi_baryi 累积的最终结果；用 V_hat_y_bar_hat 表示 $\hat{V}(\underline{y})$ ；用 FRRC 和 RRC 分别表示采用固定效应模型和随机效应模型时的合并效应 RR；用 FRRLOW、

FRRUP 及 RRLOW、RRUP 分别表示采用固定效应模型和随机效应模型时合并效应 RR 的 95% CI 的下限和上限。

【输出结果及解释】

number_of_study	Q	DF	P	I^2	FRRC
7	9.92947	6	0.12765	39.5738	0.91448

FRRLOW	FRRUP	RRC	RRLOW	RRUP
0.86653	0.96509	0.89326	0.80093	0.99624

异质性检验的结果为 $Q=\chi^2=9.92947$, $df=7-1=6$, $P=0.12765>0.05$, $I^2=39.5738\%<50\%$, 因此认为各研究同质性较好, 采用固定效应模型进行综合分析是合适的。结果表明: 平均效应尺度(RR)约为 0.91, 其 95% 置信区间为 0.87 ~ 0.97。由于置信区间位于 1 的左侧, 所以试验组与对照组的死亡危险率之间的差别具有统计学意义, 7 项研究的综合结果显示, 使用阿司匹林预防心肌梗死后死亡具有一定的效果。使用阿司匹林预防后, 可以使死亡的危险降低约 9 个百分点。

2.4.4　用 SAS 软件对例 2-3 的资料进行 CMH χ^2 检验

对例 2-3 进行 CMH 校正的 χ^2 检验, SAS 程序(程序名为 TJFX2_3. SAS)如下:

```
data tjfx2_3;
    do a=1 to 2;
    do b=1 to 2;
    do c=1 to 3;
input f @@ ;
output;
end; end; end;
cards;
52   7   2
34   8   1
23  19   4
18  11   3
;
run;
ods html;
proc freq;
    weight f;
    tables b* a* c/cmh;
run;
ods html close;
```

【程序说明】 a 表示疾病类型, a=1 表示甲, a=2 表示乙; b 表示患者性别, b=1 表示男性, b=2 表示女性; c 表示药物种类, c=1 表示中药, c=2 表示西药, c=3 表示复方药。调用 freq 过程进行 CMH 卡方检验, tables 语句中, 将调整因素放置前面, 对于该程序, 即控制住变量 b, 考察变量 a 与 c 之间的关系。

【输出结果及解释】

FREQ 过程

a * c 的汇总统计量

b 的控制

Cochran-Mantel-Haenszel 统计量(基于表得分)

统计量	对立假设	自由度	值	概率
1	非零相关	1	17.0349	<0.0001

2	行均值得分差值	1	17.0349	<0.0001
3	一般关联	2	19.0157	<0.0001

统计结论：由上述结果可知，在控制了 b 因素（即“患者性别”）后，一般关联值为 χ^2_{CMH} = 19.0157，P <0.0001，表明不同疾病患者所用的药物种类频数构成之间的差别有统计学意义。

专业结论：可以认为患甲病的患者所用的药物种类频数构成与患乙病的患者所用的药物种类频数构成是不同的。

2.4.5　用 SAS 软件对例 2-4 的资料进行 CMH 校正的秩和检验

对例 2-4 进行 CMH 校正的秩和检验，SAS 程序（程序名为 TJFX2_4. SAS）如下。

```
data tjfx2_4;
do a =1 to 2;
do b =1 to 3;
do c =1 to 3;
    input f @@ ;
    output;
end; end; end;
cards;
10  26  10   8   22  5
5   13   2   20   6   18
22  10  15  36   1   1
;
run;
ods html;
proc freq;
    weight f;
    tables a* b* c/nopercent
    nocol chisq scores =rank cmh2;
run;
ods html close;
```

【程序说明】 a 代表性别，a =1、2 分别代表男性、女性；b 代表年龄段，b =1 代表 30 ~ 39，b =2 代表 40 ~49，b =3 代表≥50；c 代表近视度数，c =1、2、3 分别代表 <3.00 D、3.00 ~ 6.00 D、>6.00 D。

freq 过程步中 tables 语句后的 nopercent 用来不显示各频数在总合计中的构成比，nocol 用来不显示各频数在列合计中的构成比。若读者认为不需要在结果中输出各频数在行合计中的构成比，在 tables a * b * c/后加 norow 即可。

将 freq 过程步中 tables 语句后的 cmh2 换成 all，可在最后的结果中输出三个统计量，分别为非零相关、行平均得分差异、一般联系；若换成 cmh1，则在最后的结果中只输出第一个统计量，而 cmh2 本身则可输出前两个统计量。

如果结果变量的级别不是等间隔的，则可以选用修正的 Ridit 法计算，可写成以下 tables 语句：

```
tables a* b* c/nopercent nocol chisq scores =modridit cmh2;
```

【输出结果及解释】

FREQ 过程

频数 行百分比	b * c 的表 1 控制：a =1			
b	c			合计
	1	2	3	
1	10	26	10	46
	21.74	56.52	21.74	

2	8	22	5	35
	22.86	62.86	14.29	
3	5	13	2	20
	25.00	65.00	10.00	
合计	23	61	17	101

b * c 的表 1 统计量

a = 1 的控制

统计量	自由度	值	概率
卡方	4	1.6324	0.8030
似然比卡方	4	1.6785	0.7946
MH 卡方(秩得分)	1	0.8383	0.3599
Phi 系数		0.1271	
列联系数		0.1261	
Cramer V 统计量		0.0899	

WARNING: 22% 的单元格的期望计数比 5 小。

卡方可能不是有效检验。

频数　　b * c 的表 2

行百分比　　控制: a = 2

b	c			合计
	1	2	3	
1	20	6	18	44
	45.45	13.64	40.91	
2	22	10	15	47
	46.81	21.28	31.91	
3	36	1	1	38
	94.74	2.63	2.63	
合计	78	17	34	129

b * c 的表 2 统计量

a = 2 的控制

统计量	自由度	值	概率
卡方	4	28.2260	< 0.0001
似然比卡方	4	33.3804	< 0.0001
MH 卡方(秩得分)	1	19.8550	< 0.0001
Phi 系数		0.4678	
列联系数		0.4237	
Cramer V 统计量		0.3308	

b * c 的汇总统计量

a 的控制

Cochran-Mantel-Haenszel 统计量(基于秩得分)

统计量	对立假设	自由度	值	概率
1	非零相关	1	19.0141	<0.0001
2	行均值得分差值	2	22.1518	<0.0001

SAS 输出结果中首先给出的是每层的列联表及层内各种统计量的结果。列联表中，各格子内列出了频数及其在行合计中的构成比。各层中给出的统计分析结果没有太大的实际意义，读者可不予参考。

最后输出对“性别”校正了的“不同年龄段”近视人群视力水平比较的 Wilcoxon 秩和检验结果，查看统计量行平均得分差异$\chi^2_{CMH}=22.1518$，$P<0.0001$。差异有统计学意义，在统计上有理由认为不同年龄段近视人群的视力水平不相同。

2.5　高维列联表资料常规统计分析基本概念

所谓高维列联表，就是表中的定性变量的个数大于等于 3。其中，通常应该有一个定性的结果变量。众所周知，定性结果变量有三种不同的表现，即二值的、多值有序的和多值名义的。在研究目的是希望在控制其他定性原因变量的前提下，考察研究者关心的那个定性变量不同水平组的结果变量的某种率(对二值结果变量而言)或频数构成情况(对多值名义结果变量而言)或平均秩(对多值有序结果变量而言)之间的差别是否具有统计学意义，就属于常规的统计分析或差异性检验问题。

上面提及的“控制其他定性原因变量”是什么含义呢？实际上，就是将那些拟被控制的原因变量的全部水平组合视为一个具有很多水平(其水平数等于全部拟被控制的定性原因变量的水平数之乘积)的“复合型定性原因变量”，并以其各水平为“分层”的条件，对研究者关心的那个原因变量对一个定性的结果变量的影响进行“部分”分析，再将其结果进行合并计算，得出一个最终的结果。这就是所谓的 CMH 校正的χ^2 检验或 CMH 校正的秩和检验。而前文提及的加权χ^2 检验，本质上就是 CMH 校正的χ^2 检验。当然，它们之间略有区别，就是前者有三项计算结果(①独立性检验；② RR 或 OR；③检验 RR 或 OR 是否等于 1)，而后者只有后两项计算结果。

2.6　高维列联表资料常规统计分析原理概述

2.6.1　加权χ^2 检验

加权χ^2 检验主要用于分析 $q\times2\times2$ 形式的三维列联表资料。根据研究的目的，若不想用复杂的对数线性模型或 logistic 回归模型来分析三维列联表资料，且资料又不适合采用简单合并方式处理，就可采用加权χ^2 检验。该检验方法无法回答被合并掉的那个原因变量对结果变量的影响作用有多大，只是评价另一个原因变量对结果变量的影响时将其对结果变量的影响扣除掉。

将第 h 层四格表所对应的观察频数 n_{h11}、n_{h12}、n_{h21}、n_{h22} 分别记为 a、b、c、d，合计频数 n_h 记为 n，则加权χ^2 检验公式如下：

$$\chi^2_{加权} = \frac{\left\{\sum\left[(ad-bc)/n\right]_h / \sum\left[(a+b)(c+d)/n\right]_h\right\}^2}{\sum\left[(a+b)(c+d)(a+c)(b+d)/n^3\right]_h / \left\{\sum\left[(a+b)(c+d)/n\right]_h\right\}^2} \quad (2\text{-}1)$$

该统计量服从自由度为 1 的卡方分布。合并的 OR 值和/或 RR 值计算公式为

$$OR = \sum(ad/n)_h / \sum(bc/n)_h \quad (2\text{-}2)$$

$$RR = \sum\left[a(c+d)/n\right]_h / \sum\left[(a+b)c/n\right]_h \quad (2\text{-}3)$$

对 OR 值和 RR 值进行统计学检验的计算公式为

$$\chi^2_{MH} = \left\{\sum\left[(ad-bc)/n\right]_h\right\}^2 / \sum\left[(a+b)(c+d)(a+c)(b+d)/(n-1)/n^2\right]_h \quad (2\text{-}4)$$

该统计量服从自由度为 1 的卡方分布。

2.6.2　CMH 统计分析

CMH 统计分析是在 MH 统计分析方法的基础上发展并提出来的，现在统称为扩展的 MH 卡方统计量，也统称之为 MH 检验。用于分层分析(即控制分层因素后)对 2×2、$R\times2$、$2\times R$、$R\times C$ 列联表资料的统计处理。换言之，CMH 检验可以控制一个或多个原因变量，考察被保留的一个原因变量 X 与结果变量 Y 之间的关联或 X 处于不同水平条件下 Y 的平均秩之间的差别是否具有统计学意义。根据 $R\times C$ 列联表网格中行变量与列变量的属性不同，在 SAS 中，CMH 输出结果给出了三种检验统计量。

第一种：非零相关统计量。适用于被考察的原因变量 X 和结果变量 Y 都是有序变量的情况，可以获得类似秩相关分析的检验结果，但没有给出秩相关系数，也未体现出相关的方向。该统计量服从自由度为 1 的卡方分布。条件如不满足则计算该统计量没有意义。

第二种：行平均分差异统计量。适用于仅结果变量为有序变量的情况。计算该统计量，可以获得 CMH 校正的秩和检验结果。

第三种：一般关联统计量。适用于原因变量 X 与结果变量 Y 均为无序变量或原因变量 X 是有序变量而结果变量 Y 为无序变量的情况。

(1)当结果变量为二值变量时，用 CMH 检验得出的以上三个统计量数值相等，CMH 校正的 χ^2 统计量即为第三种统计量。

第 h 层频数 n_{h11} 对应的的理论频数 m_h 及方差 v_h 为

$$m_h = \frac{n_{h1\cdot}n_{h2\cdot}}{n_h},\ v_h = \frac{n_{h1\cdot}n_{h2\cdot}n_{h\cdot1}n_{h\cdot2}}{n_h^2(n_h+1)}$$

CMH 校正的 χ^2 统计量的计算公式为

$$\chi^2_{CMH} = \frac{\left(\sum_h n_{h11} - \sum_h m_h\right)^2}{\sum_h v_h} \quad (2\text{-}5)$$

该统计量服从自由度为 1 的卡方分布。

(2)当结果变量为多值有序变量时，根据原因变量为多值有序变量或名义变量，分别选择其中的非零相关统计量或行平均得分统计量。

在计算非零相关统计量时，列的评分阵 $\boldsymbol{C}_h$ 是 $1\times C$ 阵，行的评分阵 $\boldsymbol{R}_h$ 是 $1\times R$ 阵，行与列的评分由 FREQ 过程中的 SCORES 选项指定。非零相关统计量的自由度为 1，它也被称为

Mantel-Haenszel 统计量。当行变量或列变量不是有序变量时，该统计量是没有意义的。非零相关统计量对应的备择假设为：至少在一层中，原因变量和结果变量之间存在线性相关。

在计算行平均得分统计量时，列的评分阵 $\boldsymbol{C}_h$ 是 $1 \times C$ 阵，由 SCORES 选项指定；行的评分阵 $\boldsymbol{R}_h$ 是 $(R-1) \times R$ 阵，由 FREQ 过程内部产生，即

$$\boldsymbol{R}_h = (\boldsymbol{I}_{R-1}, -\boldsymbol{J}_{R-1}) \tag{2-6}$$

式中，$\boldsymbol{I}_{R-1}$ 是秩为 $R-1$ 的单位阵；$\boldsymbol{J}_{R-1}$ 是元素均为 1 的 $(R-1) \times 1$ 的列向量。

行平均得分统计量的自由度为 $R-1$，它所对应的备择假设为：至少在一层中，R 行之间的平均得分是不同的，也就是按原因变量分为 R 个组之后，不同组别之间关于结果变量的平均得分存在差异。

(3)当结果变量为多值名义变量时，选用一般关联统计量，也就是 FREQ 过程 CMH 检验输出结果中的第三项，其自由度 $\nu = (R-1)(C-1)$。在计算该统计量时，行的评分阵 $\boldsymbol{R}_h$ 与计算行平均得分统计量的行评分矩阵相同，即 $\boldsymbol{R}_h = (\boldsymbol{I}_{R-1}, -\boldsymbol{J}_{R-1})$；列的评分阵 $\boldsymbol{C}_h$ 可以被类似地定义为

$$\boldsymbol{C}_h = (\boldsymbol{I}_{C-1}, -\boldsymbol{J}_{C-1}) \tag{2-7}$$

式中，$\boldsymbol{I}_{C-1}$ 是秩为 $C-1$ 的单位阵；$\boldsymbol{J}_{C-1}$ 是元素均为 1 的 $(C-1) \times 1$ 的列向量。在一般关联统计量中，行的评分阵 $\boldsymbol{R}_h$ 与列的评分阵 $\boldsymbol{C}_h$ 都是由 FREQ 过程内部产生的。

一般关联统计量不要求原因变量或结果变量是有序的。无论原因变量是多值有序变量还是名义变量，只要结果变量是多值名义的，都可以采用该统计量。与之相对应的原假设和备择假设分别如下。

H_0：每层中原因变量和结果变量之间不存在关联。

H_1：至少有一层，原因变量和结果变量之间存在某种关联。

当仅有一层时，该 CMH 统计量与 Pearsonχ^2 统计量的关系为

$$\chi^2_{\text{CMH}} = [(n-1)/n]\chi^2 \tag{2-8}$$

式中，n 为总例数；当有多层时，该统计量为层修正的 Pearsonχ^2 统计量。当然，相似的校正也能够通过对各层 Pearsonχ^2 统计量求和而得到，但是这种校正方法需要每层的样本含量都要足够大，而 CMH 统计量仅仅需要总的样本含量比较大。

需要说明的是，当各层的各比较组间的趋势方向一致时，CMH 方法比较有效；当各层的各比较组间的趋势方向不一致时，CMH 方法则不容易检出差别，此时应单独考察各层或采用其他方法。

2.6.3　meta 分析

单个队列研究设计四格表资料见表 2-6。

多个队列研究设计四格表资料见表 2-7。

表 2-6　队列研究设计四格表资料

暴露与否	例　数	
	发病与否：发病	未发病
暴露	a_i	b_i
不暴露	c_i	d_i

表 2-7　k 个队列研究设计四格表资料

研究编号	观测人数(发生某事件的人数)	
	阿司匹林组	安慰剂组
1	$n_{11}(n_{a1})$	$n_{12}(n_{c1})$
…	…	…
k	$n_{k1}(n_{ak})$	$n_{k2}(n_{ck})$

队列研究属于观察性研究，其效应指标有危险比或相对危险度（Risk Ratio or Relative Risk，RR）及危险差（Risk Difference，RD）。对多个符合入选条件的队列研究设计四格表资料进行 meta 分析的基本原理如下。

1. 处理效应及权重

假设已获得 k 项符合入选条件的研究，编号 $i = 1, 2, \ldots, k$。每项研究均有 1 个暴露组（记为 T）和 1 个对照组（即不暴露组，记为 C）。

有时候，队列研究设计四格表中的某个（些）频数为 0，从而给计算带来困难。为了避免这种情况的发生，通常的做法是对 4 个频数中的每一个都加上一个很小的值，如 0.005 或 0.0025。

可用表 2-6 中的 4 个频数来估计暴露组和对照组受试者发生研究者所关心的某事件的发生率，公式如下：

$$\hat{p}_{Ti} = a_i/(a_i + b_i) \tag{2-9}$$

$$\hat{p}_{Ci} = c_i/(c_i + d_i) \tag{2-10}$$

基于这些发生率，可以进一步计算效应指标 RR_i 及 RD_i 的值，即

$$\hat{RR} = p_{Ti}/p_{Ci} \tag{2-11}$$

$$\hat{RD} = p_{Ti} - p_{Ci} \tag{2-12}$$

进行统计分析时，通常对 RR_i 作对数变换，将变换后的变量作为处理效应。这是因为对于小样本研究，作对数变换后的变量的分布与正态分布更为接近。对 RR_i 作对数变换后的变量，其所对应的总体方差的估计值为

$$\hat{V}[\ln(\hat{RR}_i)] = \frac{1}{a_i} - \frac{1}{a_i + b_i} + \frac{1}{c_i} - \frac{1}{c_i + d_i} \tag{2-13}$$

RD_i 所对应的总体方差的估计值为

$$\hat{V}(\hat{RD}) = \frac{p_{Ti}(1 - p_{Ti})}{a_i + b_i} + \frac{p_{Ci}(1 - p_{Ci})}{c_i + d_i} \tag{2-14}$$

若用 y_i 来统一表示 $\ln(RR_i)$ 或 RD_i 中的任何一个，用 $\hat{y}_i$ 表示 y_i 的估计值，则 y_i 基于正态分布的置信区间为

$$\hat{y}_i \pm z_{1-\alpha/2}\sqrt{\hat{V}(\hat{y}_i)} \tag{2-15}$$

当效应指标为 RR_i 时，式(2-15)的置信区间为作对数变换后的新变量的置信区间，可通过再变换得到原始变量的置信区间。

各项研究效应指标的权重定义如下：

$$v_i = \hat{V}(\hat{y}_i) \tag{2-16}$$

$$w_i = 1/v_i \tag{2-17}$$

2. 假设检验

人们已经提出了几种用于检验所感兴趣的不同假设的检验方法。

整体零假设的检验，检验假设为 $H_0: y_i = 0$，$i = 1, 2, \cdots, k$，即所有的处理效应均为零。

对于该假设，有两种统计学检验方法，一种叫作不定向检验，其对立假设为“至少有 1 项研究的处理效应 $y_i \neq 0$”，通过比较统计量

$$X_{ND} = \sum_{i=1}^{k} w_i i^2 \tag{2-18}$$

与自由度为 k 的 χ^2 分布的临界值 χ^2_k 的大小来实现。

另一种叫作定向检验，其对立假设为“各项研究的处理效应均为相同的非零值 y，即 $y_i = y \neq 0$”，通过比较统计量

$$X_D = \left(\sum_{i=1}^{k} w_i \hat{y}_i\right)^2 / \sum_{i=1}^{k} w_i \tag{2-19}$$

与自由度为 1 的 χ^2 分布的临界值 χ_1^2 的大小来实现。

当整体零假设被拒绝后，下一步就是检验所有各项研究的处理效应是否相等，也就是检验各项研究的处理效应是否同质。检验零假设为 $H_0: y_i = y$，$i = 1, 2, \cdots, k$，对立假设为“至少有 1 项研究的效应与其他研究不同，即各项研究的效应是不同质的”。

对异质性的检验采用 Cochran's Q 检验，统计量为

$$Q = \sum_{i=1}^{k} w_i (\hat{y}_i - \hat{y})^2 \tag{2-20}$$

式中，

$$\hat{y} = \left(\sum_{i=1}^{k} w_i \hat{y}_i\right) / \sum_{i=1}^{k} w_i \tag{2-21}$$

该检验通过比较检验统计量 Q 与自由度为 $k-1$ 的 χ^2 分布的临界值 χ_{k-1}^2 的大小来实现。

关于 meta 分析中研究间异质性的度量，Julian P. T. Higgins 等提出了一个新的更好的统计量 I^2，由下式计算得到：

$$I^2 = \frac{Q - (k-1)}{Q} \times 100\% \tag{2-22}$$

式中，Q 为前面的异质性检验 χ^2 检验统计量；k 为入选研究的个数。

统计量 I^2 度量了各研究间由抽样误差以外的因素引起的异质性占总的异质性的比例。根据Julian P. T. Higgins 等所描述的计算方法，当计算出的 I^2 为负值时，取 $I^2 = 0$，因此，I^2 的取值为 0 ~ 100% 之间。当 $I^2 > 50\%$ 时，即认为各研究间是异质的。

完成异质性检验后，进一步的工作就是选择正确的模型来构造合并效应的置信区间。

不管是通过统计检验还是从别的方面考虑，如果认为各研究的效应是相等的(同质的)，则可以选择固定效应模型来构造合并效应的置信区间；反之，如果各研究的效应是异质的，则应选择随机效应模型构造合并效应的置信区间。

当选择固定效应模型时，按照下式计算 y 的可信区间：

$$\hat{y} \pm z_{1-\alpha/2} \sqrt{\hat{V}(\hat{y})} \tag{2-23}$$

式中，$z_{1-\alpha/2}$ 为标准正态分布的 $(1-\alpha/2) \times 100\%$ 百分位点，且

$$\hat{y} = \left(\sum_{i=1}^{k} w_i \hat{y}_i\right) / \left(\sum_{i=1}^{k} w_i\right) \tag{2-24}$$

$$\hat{V}(\hat{y}) = \left(\sum_{i=1}^{k} w_i\right)^{-1} \tag{2-25}$$

当选择随机效应模型时，按照下式计算 y 的置信区间：

$$\hat{\bar{y}} \pm z_{1-\alpha/2} \sqrt{\hat{V}(\hat{\bar{y}})} \tag{2-26}$$

式中，$z_{1-\alpha/2}$ 为标准正态分布的 $(1-\alpha/2) \times 100\%$ 百分位点，且

$$\hat{\bar{y}} = \left(\sum_{i=1}^{k} \bar{w}_i \hat{y}_i\right) / \sum_{i=1}^{k} \bar{w}_i \tag{2-27}$$

$$\hat{V}(\hat{\bar{y}}) = \left(\sum_{i=1}^{k} \bar{w}_i\right)^{-1} \tag{2-28}$$

$$\bar{w}_i = \left(\frac{1}{w_i} + \hat{\tau}^2\right)^{-1} \tag{2-29}$$

$$\hat{\tau}^2 = \begin{cases} \dfrac{Q-(k-1)}{U} & Q > k-1 \\ 0 & Q \leqslant k-1 \end{cases} \tag{2-30}$$

式中的 Q 由式(2-20)计算。

$$U = (k-1)\left(\bar{w} - \frac{s_w^2}{k\bar{w}}\right) \tag{2-31}$$

$$s_w^2 = \frac{1}{k-1}\left(\sum_{i=1}^{k} w_i^2 - k\bar{w}^2\right) \tag{2-32}$$

$$\bar{w} = \frac{1}{k}\left(\sum_{i=1}^{k} w_i\right) \tag{2-33}$$

式中的 w_i 由式(2-17)计算。

2.7　本章小结

通常，列联表可分为 2×2 表(4 类)、$R \times C$ 表(5 类)、高维列联表(3 类)和具有重复测量因素的列联表。本章结合实例介绍了高维列联表资料(结果变量为二值、多值有序和多值名义)的常规统计分析方法、SAS 程序、结果解释和基本原理。分析定性资料时，首先应正确判断资料所对应的列联表类型；其次根据不同的分析目的，并结合统计分析方法的应用条件，选择合适的分析方法。

参考文献

[1] 胡良平. 医学统计学—运用三型理论分析定量与定性资料[M]. 北京:人民军医出版社, 2009:325-393.

[2] 胡良平. 心血管病科研设计与统计分析[M]. 北京:人民军医出版社, 2010:133-137.

[3] 胡良平. 面向问题的统计学—(2)多因素设计与线性模型分析[M]. 北京:人民卫生出版社, 2012:547-566.

[4] 方积乾. 医学统计学与电脑实验[M]. 3 版. 上海:上海科学技术出版社, 2006:459-468.

[5] 方积乾. 生物医学研究的统计方法[M]. 北京:高等教育出版社, 2007: 502-527.

[6] Julian P T Higgins, Simon G Thompson, Jonathan J Deeks, et al. Measuring inconsistency in Meta-analyses [M]. BMJ, 2003(327):557-560.

[7] Xiao-Hua Zhou, Nancy A Obuchowski, Donna K McClish. 诊断医学统计学[M]. 宇传华,译. 北京:人民卫生出版社, 2005:2-293.

[8] SAS Institute Inc. SAS/STAT 9.2 User's Guide. 2008:1675-1829.

[9] 胡良平. SAS 统计分析教程[M]. 北京:电子工业出版社, 2010:160-182.