دانشگاه خلیج فارس

دانشکده علوم پایه

مقطع کارشناسی

پروژه کارشناسی

موضوع پروژه :

رگرسیون چند متغیره
و تحلیل های آن در نرم افزار SAS

استاد مشاور :

دکتر محمد اسماعیل دهقان منفرد

استاد داور:

دکتر فضل الله لک

محققین :

سید مرتضی نجیبی – حسن مزارعی

تابستان ۸۷

بسم الله الرحمن الرحيم

تقدیر و تشکر :

در اینجا جا دارد از زحمات فراوان اساتید محترم گروه آمار و ریاضی دانشگاه خلیج فارس از جمله جناب آقای دکتر محمد اسماعیل دهقان منفرد،دکتر فضل الله لک ، دکتر محمود افشاری و آقای مراد علیزاده که ما را در این مهم یاری کردند،تشکر و قدردانی کنیم.

فهرست

فصل اول

پیشگفتار

فصل اول :پیشگفتار

۱– مقدمه

۱–۱– اهداف پروژه

کاربران آماری علوم بیولوژی ، علوم اجتماعی،روانشناسی و فیزیک معمولا اندازه گیری هایی روی چندین متغیر را انجام داده و با جمع آوری داده های مربوطه سعی در کشف ارتباط بین این داده ها را دارند.

با توجه به اینکه رابطه های چند متغیره در دوره های کارشناسی در دانشگاه های کشور کمتر مورد توجه اساتید و دانشجویان قرار می گیرد ، لذا ما برای یادگیری بیشتر در این زمینه بر روی رگرسیون چند متغیره مطالعاتی انجام دادیم.هر چند تحلیل داده های یک متغیره نسبت به چند متغیره به سادگی انجام می شود و کاربرد بسیار زیادی در زندگی روزمره ما دارد ولی وقتی با چند متغیر سر و کار داشته باشیم بررسی و تحلیل آنها بسیار جالب و جذاب و در عین حال نسبتا پیچیده است .

هدف ما از انجام این پروژه ارائه مفاهیم و روشهای تحلیل رگرسیونی چند متغیره در سطح دانشجویان کارشناسی و آشنایی بیشتر خواننده با نرم افزار های آماری مربوط به تحلیل های چند متغیره می باشد.در این پروژه ابتدا مقدمه ای از رگرسیون چند متغیره را در فصل دوم ارئه داده و برای یادگیری تحلیل های کاربردی در این زمینه با ارائه داده های واقعی

به طریقه تحلیل و تفسیر آنها به کمک نرم افزار آمـاری SAS 9.1 در فـصل هـای بعـد مـی پردازیم.

۱-۲- خلاصه ای از نرم افزار SAS

SAS مخفف Statistical Analysis System ، یک نرم افزار پرقـدرت آمـاری است. این نرم افزار بیشتر روشهای رایج آماری را پوشش می دهد و کاربر می‌تواند فرمانهـای خاصی را که متناسب با نیازهایش باشد انتخاب نموده و به‌کار گیرد.

گزینه ها و زیرگزینه‌های متعددی در SAS وجود دارد که کاربر را قادر مـی‌سـازد تا تجزیه و تحلیلهای آماری و غیر آماری را در سطوح مختلفی از نظـر پیچیـدگی و جزئیـات انجام دهد . انواع تجزیه و تحلیلهای سریهای زمانی ،انواع مدلهای خطی و غیر خطی ،روشهای چند متغیره پیوسته و گسسته ،کنترل کیفیت ،آمار توصیفی، انواع تحلیلهای گرافیکی و نموداری ،انواع تحلیلهای ماتریسی و ... را می‌توان توسط این نـرم افـزار انجـام داد. بـر خـلاف نـرم افزارهایی مانند Minitab و Spss که ارتباط کاربر با نرم افزار از طریق پنجره ها و جعبه های محاوره‌ای صورت می‌گیرد ،SAS احتیاج به برنامه نویسی دارد. پس از نـسخه تحـت Dos ، تاکنون نسخه‌های ۸٫۷ و ۹٫۱ از این نرم افزار که تحت ویندوز هستند نیز وارد بازار شده است . البته درنسخه‌های ۸٫۷ و ۹٫۱ برخی دستورات به صورت پنجره ارائه شـده اسـت امـا روش اصلی در کار با SAS برنامه نویسی است.

از مهمترین مزایای این نرم افزار قابل استفاده بودن آن برای همـه کـاربران اعـم از مبتدی و پیشرفته است . از محاسن دیگر این نرم افزار امکان برنامـه نویـسی بـرای تحلیلهـای

پیشرفته آماری و علی‌الخصوص برنامه نویسی جهت تحلیلهای ماتریسی است . ولـی SAS در برخی زمینه ها از جمله درزمینه گرافیکی از سایر نرم‌افزارها نظیرSplus ضعیف تر است.

در کل نرم افزار SAS یکی از قویترین نرم افزارهای آماری بـه حـساب مـی‌آیـد و در اکثـر کشورها به طور گسترده مورد استفاده قرار می‌گیرد ، به طوری که در برخی کشورهای پیشرفته از لحاظ آمار از جمله در آمریکا ، آشنایی و توان استفاده از این نـرم افـزار بـرای یـک آمـار شناس از اهمیت ویژه ای برخوردار است.

علاقه‌مندان به یادگیری SAS می‌توانند از Help این نرم افزار استفاده کنند کـه بـه صورتی ساده،روان و بدون استفاده زیاد از لغات تخصصی ، به طور کامل ایـن نـرم افـزار را معرفی نموده و آموزش می‌دهد.

۱–۳– روند مطالعات و اقدامات لازم برای انجام پروژه

با توجه به اینکه ما دانشجویان کارشناسی رشته آمـار تنهـا رگرسـیون چنـد گانـه (رگرسیون خطی یک متغیره) و تحلیل های چند متغیره را دراین دوره در دروس رگرسیون و تحلیل های چند متغیره پیوسته مطالعه می کنیم ، لذا ما می بایست ابتدا بـه مفهـوم رگرسـیون چند متغیره در منابع مختلف بپردازیم. برای این منظور منبع فارسی خود را کتاب تحلیل آماری چند متغیری کاربردی تالیف ریچارد آ.جانسون و دین . دبلیو.ویچرن ترجمـه دکتـر حـسینعلی نیرومند قرار داده و جهت تکمیل اطلاعـات لازم در ایـن زمینـه بـا اسـتفاده از کتـب لاتـین Alvin C.Rencher – Method of Multivariate Analysis و Applied Multivariat Analysis, Neil H به تکمیل اطلاعات خود پرداختیم. با توجه بـه پیچیـدگی

های موضوع ، گسترده بودن آن و همچنین فقدان اطلاعات کافی در این زمینه، ممکن است علی رغم تلاش های فراوان ما ،کم و کاستی هایی در این پروژه موجود باشد.

پس از کسب اطلاعات اولیه بر آن شدیم تا مثال هایی واقعی از کاربرد رگرسیون چند متغیره در جامعه آورده و تحلیل های لازم را با کمک نرم افزار SAS ارائه دهیم. برای این منظور ار منبع ۳ استفاده کرده و پس از یادگیری برنامه های مربوطه و تغییرات جزئی در آنها به رسم نمودار های مربوطه و تحلیل آنها در حد توان پرداختیم.

فصل دوم

رگرسیون چند متغیره

فصل دوم : رگرسیون چند متغیره

۲-۱- آشنایی اولیه[1]

اهدافی که به بیان یک پدیده اجتماعی یا فیزیکی مربوط می شود بایستی مشخص شده ، و سپس با گردآوری داده ها ، آزمون و تحلیل شوند. با توجه به اینکه عوامل و متغیر های زیادی بر یک پدیده تاثیر دارند و داده ها اندازه های همزمانی را در مورد چند متغیر شامل می شوند، تحلیل این داده ها روشی را در علم آمار پایه ریزی می کند که به روش تحلیل چند متغیره نامیده می شود.

لزوم درک و بررسی روابط بین بسیاری از متغیر ها، تحلیل چند متغیره را مشکل می سازد. اغلب فکر بشری در انبوهی از داده ها غوطه ور می شود و لذا تحلیل های چند متغیره پیچیدگی های بیشتری را شامل می شود.

بطور فزاینده ای معلوم می شود که بسیاری از روش های چند متغیره بر اساس الگوی احتمال مورد بررسی ، که به توزیع نرمال چند متغیره موسوم است ، قرار دارد. روشهای دیگر دارای طبیعت خاصی هستند و با استدلالهای منطقی و یا با عقل سلیم تایید می شوند. پیشرفتهای اخیر در فن آوری رایانه همراه با توسعه بسته های نرم افزاری آماری پیشرفته ، مرحله اجرا را ساده تر می کند .

---

[1] - منبع یک

۲-۱-۱- اهداف بررسی های علمی روش های چند متغیره

اهداف بررسی های علمی که برای آن ها روش های چند متغیره بطور خیلی طبیعی بکار می روند ، شامل موارد زیر است :

أ- **کاهش داده ها یا سهولت ساختاری** (پدیده مورد مطالعه ، بدون این که اطلاعات با ارزشی را از دست دهیم ، تا جایی که ممکن است ساده ارائه می شود ، امیدواریم که این امر تعبیر و تفسیر را ساده تر کند)

ب- **جور کردن و دسته بندی کردن** (دسته های اشیا یا متغیر های (مشابه) بر مبنای خصیصه های اندازه گیری شده ، ایجاد می شوند.)

ج- **بررسی وابستگی میان متغیرها** (طبیعت روابط میان متغیر ها مورد علاقه است. آیا بطور طبیعی تمام متغیر ها مستقل اند یا اینکه یک یا چند متغیر به سایرین وابسته اند ؟اگر چنین است میزان وابستگی آنها تا چه میزان می باشد.)

د- **ساختن و آزمون کردن فرضها** (فرض های آماری خاصی که بر حسب پارامتر های جامعه های چند متغیره فرمول بندی می شوند ، را آزمون می کنیم .این را می توان با فرض های معتبر یا تقویت کردن عقاید پیشین انجام داد)

۲-۱-۲- کاربرد های روش های چند متغیره

روش های آماری بخش لاینفک تحقیقات علمی است ،در نتیجه کاربرد آنهـا بـسیار زیاد است . بویژه الگوهای چندمتغیره به وفور در مسائلی که در علوم فیزیکـی ، روانـشناسی، پزشکی، جامعه شناسی ، اقتصاد بازرگانی ، علوم تربیتی ، زیست شناسی ، مطالعات مربوط بـه یک محیط ،هواشناسی و زمین شناسی پیش می آید ، کاربرد دارد .

۲-۲- تعاریف اولیه

با توجه به آشنایی شما خواننده محترم با رگرسیون خطی ساده و خطی چند گانه در درس رگرسیون ما ابتدا تعاریف مختصری از رگرسیون خطی ساده و چندگانـه را ارائـه داده و سپس آن را به رگرسیون چند متغیره خطی تعمیم می دهیم.

۲-۲-۱- رگرسیون خطی ساده

همانطور که می دانید در رگرسیون خطی ساده هدف این است که رابطه بـین یـک متغیر مستقل و یک متغیر وابسته را بررسی کنیم . در اینجا مدل رگرسیونی را بصورت

رابطه (۱)
$$\mathrm{E}(Y|X_1 = x_1) = \beta_0 + \beta_1 x_1$$

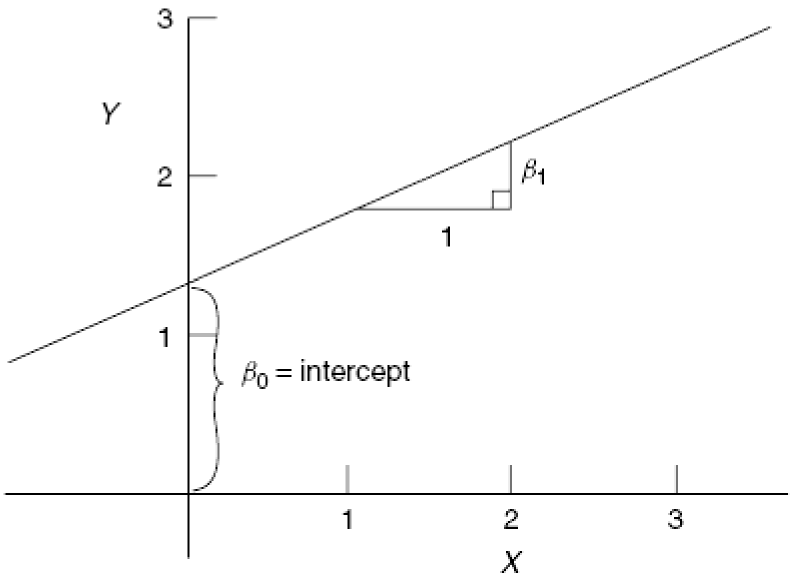
در نظر می گیریم که در آن

$$\mathrm{Var}(Y|X=x)=\sigma^2$$
رابطه (۲)

و همچنین اگر مدل را بصورت $y_i=\beta_0+\beta_1 x_i+\varepsilon_i$ در نظر بگیریم داریم :

$\varepsilon_i \sim N(0,\sigma^2)$ می باشد .

در زیر نمایی از رگرسیون خطی ساده بر روی شکل نمایش داده شده است :



شکل **1**

که در این مدل با روش حداقل مربعات برآورد پارامتر های مدل ،بصورت زیر بدست آمد:

$$\hat{y}_i=\widehat{\mathrm{E}}(Y|X=x_i)=\hat{\beta}_0+\hat{\beta}_1 x_i$$
رابطه (۳)

$$\hat{\beta}_1=\frac{SXY}{SXX}$$
رابطه (۴)

$$\hat{\beta}_0=\overline{y}-\hat{\beta}_1\overline{x}$$
رابطه (۵)

$$\hat{\sigma}^2=\frac{SSE}{n-2}$$
رابطه (۶)

۲-۲-۲- رگرسیون خطی چندگانه

همانطور که می دانید در رگرسیون چند گانه متغیر هـای مـستقل(توضـیحی) بـه P متغیر تعمیم می یابد و مدل رگرسیونی مربوط به آن بصورت :

$$\mathrm{E}(Y|X) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$
رابطه (۷)

$$\mathrm{Var}(Y|X) = \sigma^2$$
رابطه (۸)

این مدل را می توان بصورت ماتریسی نیز نمایش داد:

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$
رابطه (۹)

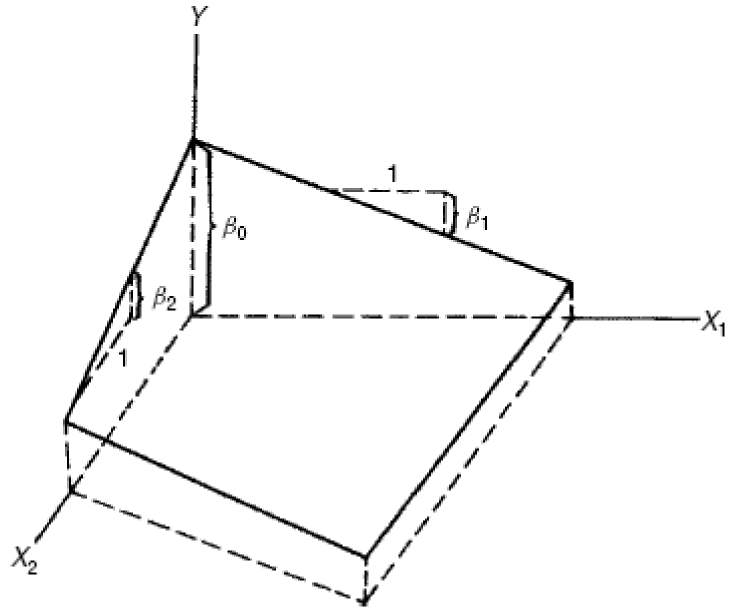
$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}$$
رابطه (۱۰)

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \quad \text{and} \quad \mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$
رابطه (۱۱)

$$\begin{aligned}\mathbf{E}(Y|X = \mathbf{x}_i) &= \mathbf{x}_i'\boldsymbol{\beta} \\ &= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}\end{aligned}$$
رابطه (۱۲)

پس مدل رگرسیونی بصورت $\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}$ در می آید .

در زیر نمایی سه بعدی از حالتی که p=2 می باشد آمده است :



شکل 2

و همچنین برآورد پارامتر های مدل با روش حداقل مربعات بصورت زیر بدست می آید :

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$ رابطه (۱۳)

و همچنین

$$\hat{\sigma}^2 = \frac{\mathbf{SSE}}{\mathbf{n} - (\mathbf{p} + 1)}$$ رابطه (۱۴)

همانطور که دیدیم در رگرسیون خطی ساده و چند گانه یک متغیر وابسته داشتیم و یک یا چند متغیر توضیحی ، که تحلیل های ساده ای را نسبت به رگرسیون چند متغیره دارد.

توجه کنید که منظور از اصطلاح رگرسیون خطی این است که میانگین تابعی خطی از پـارامتر های نامعلوم $\beta_0, \beta_1, \beta_2, ..., \beta_q$ می باشد.

یک سری فرضیات بر روی مدل های رگرسیونی گذاشته مـی شـود کـه مـا بـرای جلوگیری از تکرار در مدل های خطی ساده و چندگانه از روش هـای آزمـون ایـن فرضـیات خودداری کردیم بعنوان مثال در رگرسیون چندگانه داریم :

1. $E(e_i)=0$
2. $Var(e_i)= \sigma 2$
3. $Cov(e_i,e_j)=0$
4. $Cov(x_i,x_j)=0$

ولی در بخش چند متغیره بطور کامل این شرایط را توضیح و طریقه بررسی درست بودن آنهـا را با مثال نشان می دهیم.

۲-۲-۳- رگرسیون چند متغیره

می توان گفت الگوی رگرسیونی چند متغیره تعمیم رگرسیون چندگانه یک متغیره است.در رگرسیون چندگانه یک متغیره ما تعدادی متغیر مستقل در دسترس داریم که می خواهیم با استفاده از آنها میزان اثر متغیر های مستقل را بر یک متغیر وابسته (متغیر پاسخ) بررسی کنیم.

در رگرسیون چند متغیره چندین متغیر مستقل و چندین متغیر وابسته داریم و هدف تحلیل اثرات متغیر های مستقل بر چند متغیر وابسته (پاسخ) می باشد. ابتدا در این فصل میزان تاثیر پذیری متغیر های وابسته را بررسی کرده و سپس به کمک متغیر های مستقل مقادیر متغیر های وابسته را پیش بینی می کنیم.برای مثال می خواهیم میزان اثر بخشی شدت وزش باد بر مهارت خلبانی از جمله سرعت عکس العمل،میزان دید،مهارت ریاضیات و علوم و ... بررسی کنیم[۲].در این مثال شدت وزش باد بعنوان متغیر مستقل و متغیر های دیگر بعنوان متغیر وابسته می باشند.

متغیر های مستقل به دو نوع متغیر های ثابت و تصادفی تقسیم می شوند.متغیر های تصادفی تحت کنترل محقق نمی باشند،اما در انجام کار های رگرسیونی معمولا متغیر های توضیحی را ثابت فرض می کنیم.

[2] منبع ۲

۲-۳- الگو های رگرسیون خطی چند متغیره

۲-۳-۱- الگوی رگرسیونی خطی چند متغیره با متغیر های توضیحی ثابت

فرض کنید $x_1, x_2, \ldots, x_q$ متغیر های مستقل ثابت بوده و $y_1, y_2, y_3, \ldots, y_P$ ، P متغیر وابسته باشند.که می توان آن را با مدل های رگرسیونی زیر نمایش داد :

$$\begin{aligned} \mathbf{Y_1} &= \beta_{01} + \beta_{11}\mathbf{X_1} + \ldots + \beta_{q1}\mathbf{X_q} + \varepsilon_1 \\ \mathbf{Y_2} &= \beta_{02} + \beta_{12}\mathbf{X_1} + \ldots + \beta_{q2}\mathbf{X_q} + \varepsilon_2 \\ &. \\ &. \\ &. \\ \mathbf{Y_p} &= \beta_{0p} + \beta_{1p}\mathbf{X_1} + \ldots + \beta_{qp}\mathbf{X_q} + \varepsilon_p \end{aligned}$$

رابطه (۱۵)

که بصورت ماتریسی می توان آن را بصورت زیر نمایش داد :

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix} = \begin{pmatrix} \mathbf{y}'_1 \\ \mathbf{y}'_2 \\ \vdots \\ \mathbf{y}'_n \end{pmatrix}$$

رابطه (۱۶)

و همچنین ماتریس متغیر های مستقل بصورت :

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1q} \\ 1 & x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nq} \end{pmatrix}$$

رابطه (۱۷)

رابطه (۱۸)

$$\beta = \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ . & \cdots & \cdots & . \\ . & \cdots & \cdots & . \\ . & \cdots & \cdots & . \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{bmatrix}$$

رابطه (۱۹)

$$\Xi = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2p} \\ . & \cdots & \cdots & . \\ . & \cdots & \cdots & . \\ . & \cdots & \cdots & . \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{bmatrix}$$

پس در این حالت مدل بصورت $Y = X\beta + \Xi$ در می آید.

در اینجا β, σ_{ik} مقادیر نامعلوم می باشند.

همان طور که دیده می شود پاسخ y_i' از الگوی رگرسیون خطی زیر پیروی می کند :

$y_i' = X\beta_i' + \varepsilon_i', i=1,2,\ldots,p$ که در آن $cov(\varepsilon_i') = \sigma_{ii}$ می باشد.

بعنوان مثال اگر فرض کنیم $p = 2, q = 3$ ماتریس های آنها بصورت زیر می باشد :

$$\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \\ \vdots & \vdots \\ y_{n1} & y_{n2} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{pmatrix} \begin{pmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \\ \vdots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} \end{pmatrix}$$

و مدل برای اولین ستون y بصورت :

$$\begin{pmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{n1} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{pmatrix} \begin{pmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{21} \\ \beta_{31} \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{n1} \end{pmatrix}$$

و برای ستون دوم نیز :

$$\begin{pmatrix} y_{12} \\ y_{22} \\ \vdots \\ y_{n2} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{pmatrix} \begin{pmatrix} \beta_{02} \\ \beta_{12} \\ \beta_{22} \\ \beta_{32} \end{pmatrix} + \begin{pmatrix} \varepsilon_{12} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{n2} \end{pmatrix}$$

فرض های زیر در رگرسیون چندگانه چند متغیره وجود دارد:

۱. جمله خطای $\Xi = [\varepsilon_1, \varepsilon_2, \ldots\ldots \varepsilon_p]'$ دارای توزیع نرمال P متغیره $N_p(0, \Sigma)$ می باشد

۲. بنابراین جمله خطای $\Xi = [\varepsilon_1, \varepsilon_2, \ldots\ldots \varepsilon_p]'$ دارای $E(\varepsilon_i) = 0$ و $\mathbf{var}(\Xi) = \sum$، از این رو جملات خطاهای مرتبط با پاسخ های مختلف ، ممکن است همبسته باشند.

۳. همچنین در این مدل نیز داریم $\mathrm{cov}(\mathbf{y}_i, \mathbf{y}_j) = \mathbf{O}$ for all $i \neq j$

در فرضیه اول ماتریس واریانس ، کوواریانس شامل واریانس و کوواریانس $y_{i1}, y_{i2}, \ldots, y_{ip}$ در هر $\mathbf{y}_i$ بصورت زیر می باشد :

$$\mathrm{cov}(\mathbf{y}_i) = \mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$$

رابطه (۲۰)

در ماتریس $\mathrm{cov}(\mathbf{y}_i, \mathbf{y}_j) = \mathbf{O}$ در فرضیه ۳ مقدار کوواریانس های هر $y_{i1}, y_{i2}, \ldots, y_{ip}$ با هر $y_{j1}, y_{j2}, \ldots, y_{jp}$ بصورت :

$$\begin{pmatrix} \mathrm{cov}(y_{i1}, y_{j1}) & \mathrm{cov}(y_{i1}, y_{j2}) & \cdots & \mathrm{cov}(y_{i1}, y_{jp}) \\ \mathrm{cov}(y_{i2}, y_{j1}) & \mathrm{cov}(y_{i2}, y_{j2}) & \cdots & \mathrm{cov}(y_{i2}, y_{jp}) \\ \vdots & \vdots & & \vdots \\ \mathrm{cov}(y_{ip}, y_{j1}) & \mathrm{cov}(y_{ip}, y_{j2}) & \cdots & \mathrm{cov}(y_{ip}, y_{jp}) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

توجه داشته باشید که در مدل های رگرسیونی چند متغیره مانند یک متغیره فرض بر این است که توزیع Y ها از توزیع نرمال چند متغیره پیروی می کند،علاوه بر این در این پروژه بردار X ثابت فرض می شود.

همانطور که می دانیم در رگرسیون چند متغیره مولفه های بردار y وابسته می باشد،لذا ماتریس واریانس کوواریانس را نمی توان بصورت همگون $\Sigma = \sigma^2 \mathbf{I}_n$ در نظر گرفت.

یکی از روش های ممکن انجام روش حداقل مربعات می باشد. دو حالت عمومی دیگر که برای ماتریس واریانس کوواریانس در نظر گرفته می شود ، یکی در نظر گرفتن وزنی برای آن (یعنی $\Sigma = \sigma^2 \mathbf{V}$) می باشد،که در این حالت به آن رگرسیون وزنی گفته می شود که در آن ماتریس v معلوم و ناتکین است(WLS) . به این مدل کمترین مربعات وزنی گفته می شود و دیگری وقتی که Σ معلوم و ناتکین باشد می توان از زوش حداقل مربعات کلی (تعمیم یافته) استفاده کرد(GLS).

زمانی که Σ نامعلوم باشد می توان گفت بطور مجانبی وقتی که نمونه بزرگ باشد دارای توزیع نرمال می باشد و در این حالت از روش حداقل مربعات محتمل تعمیم یافته[3] (FGLS) و یا مدل حداقل مربعات خطی تعمیم یافته[4] (EGLS) می توان استفاده کرد.

زمانی که داده های ما شامل ناخالصی و یا توزیع آنها نرمال نباشد می توان از روش رگرسیون ربوست[5] ، رگرسیون ناپارامتری و یا روش بوت استراپ استفاده کرد. برای

---

[3] Feasible Generalized Least Squares

[4] Estimated Generalized Least Squares

[5] Robust regression

اطلاعات بیشتر به منبع ۴ و همچنین به Goldberger (1991), Neter et al. (1996) و Timm and Mieczkowski (1997, Chapter 4) مراجعه کنید.

۲-۳-۲ - برآورد کمترین توان های دوم پارامتر های مدل

برآورد کمترین توان های دوم پارامتر های مدل $\hat{B}$ مانند رگرسیون خطی یک متغیره چندگانه بصورت زیر می باشد .

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

رابطه (۲۱)

ما به این مقدار برآورد کمترین توان های دوم برای B می گوییم زیرا مقدار $\mathbf{E} = \hat{\boldsymbol{\Xi}}'\hat{\boldsymbol{\Xi}} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})$ که برابر توان دوم خطاهاست را مینیمم می کند.

برای بررسی بیشتر اگر بردار Y را بصورت $\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \ldots, \mathbf{y}_{(p)}$ افراز ستونی کنیم آنگاه بردار برآورد β بصورت زیر در می آید:

$$\begin{aligned}\hat{\mathbf{B}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \ldots, \mathbf{y}_{(p)})\\ &= [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_{(1)}, (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_{(2)}, \ldots, (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_{(p)}]\\ &= [\hat{\boldsymbol{\beta}}_{(1)}, \hat{\boldsymbol{\beta}}_{(2)}, \ldots, \hat{\boldsymbol{\beta}}_{(p)}].\end{aligned}$$

رابطه (۲۲)

بر طبق برآورد بدست آمده در رابطه ۲۱ مقدار پیش بینی برای ماتریس Y بصورت :

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\mathbf{B}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

و برآورد خطای مدل $\hat{\Xi} = \mathbf{Y} - \hat{\mathbf{Y}} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$ می باشد .

مثال ۲-۱[6]) در این مثال می خواهیم نتایج آزمایش های طراحی شده مربوط به عکس العمل های شیمیایی را بررسی کنیم.داده های مربوطه در جدول ۱ آمده است .

X_1 = دما

X_2 = غلضت

X_3 = زمان

Y_1 = درصد ماده اولیه تغییر نکرده

Y_2 = درصد ماده تبدیل شده به محصول مورد نیاز

Y_3 = درصد ماده تبدیل شده به محصول زائد

Number	y_1	y_2	y_3	x_1	x_2	x_3
1	41.5	45.9	11.2	162	23	3
2	33.8	53.3	11.2	162	23	8
3	27.7	57.5	12.7	162	30	5
4	21.7	58.8	16.0	162	30	8
5	19.9	60.6	16.2	172	25	5
6	15.0	58.0	22.6	172	25	8
7	12.2	58.6	24.5	172	30	5
8	4.3	52.4	38.0	172	30	8
9	19.3	56.9	21.3	167	27.5	6.5
10	6.4	55.4	30.8	177	27.5	6.5
11	37.6	46.9	14.7	157	27.5	6.5
12	18.0	57.3	22.2	167	32.5	6.5
13	26.3	55.0	18.3	167	22.5	6.5
14	9.9	58.9	28.0	167	27.5	9.5
15	25.0	50.3	22.1	167	27.5	3.5
16	14.1	61.1	23.0	177	20	6.5
17	15.2	62.9	20.7	177	20	6.5
18	15.9	60.0	22.1	160	34	7.5
19	19.6	60.6	19.3	160	34	7.5

جدول **1**

[6] منبع ۲

مثال با استفاده از فرمول بالا برآورد B بصورت زیر است :

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$
$$= \begin{pmatrix} 332.11 & -26.04 & -164.08 \\ -1.55 & .40 & .91 \\ -1.42 & .29 & .90 \\ -2.24 & 1.03 & 1.15 \end{pmatrix}$$

توجه کنید که ستون اول B متشکل از برآوردهای $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ برای رگرسیون y_1 بـر روی x_1, x_2, x_3 می باشد و ســتون دوم متــشکل از برآوردهــای $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ بـرای رگرسیون y_2 بر روی x_1, x_2, x_3 می باشد و ... .

۲-۳-۳- خواص برآورد های کمترین توان های دوم

توجــه کنیــد کــه بــرآورد حــداقل مربعــات ربطــی بــه فرضــیات مــدل ماننــد $\mathbf{E}(\varepsilon_i) = \mathbf{0}$ و $\mathbf{var}(\varepsilon) = \underset{\sim}{\sum}$ ، ندارد.

۱- برآورد کمترین توان های دوم نااریب اند یعنی $E(\hat{\mathbf{B}}) = \mathbf{B}$

۲- بر طبق قضیه گاوس مارکف این برآورد در بین تمام برآورد گرهــای نااریــب کمترین واریانس را دارا می باشد.

۳- تمــام $\hat{\beta}_{jk}$ هــا در مــاتریس $\mathbf{B}$ بــا یکــدیگر همبــسته مــی باشــند زیــرا $x_1, x_2, \ldots, x_q$ وابسته می باشند و بنابراین $\hat{\beta}$ ســتون هــای مختلــف مربوط به $\hat{\mathbf{B}}$ با یکدیگر همبسته می باشند .بعبارت دیگر $\hat{\beta}$ در هر ســتون بــا $\hat{\beta}$ در ستون دیگر وابسته است زیرا $y_1, y_2, \ldots, y_p$ همبسته اند.

بدلیل خاصیت سوم گفته شده در بالا (همبستگی بین $\hat{\beta}$ ها) در آزمون فرض های مربوط بـه β نمی توان از آزمون F که در رگرسیون چند گانه بود،استفاده کرد زیرا همه ی آزمون هـای F را نمی توان بطور همزمان در سطح α نگاه داشت.

۲-۳-۴- برآورد $\sum$

یک برآورد ناریب $\mathrm{cov}(\mathbf{y}_i) = \mathbf{\Sigma}$ برابر :

$$\mathbf{S}_e = \frac{\mathbf{E}}{n-q-1} = \frac{(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})}{n-q-1}$$

$$= \frac{\mathbf{Y}'\mathbf{Y}-\hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}}{n-q-1}.$$

رابطه (۲۳)

یعنی : $E(\mathbf{S}_e) = \mathbf{\Sigma}$

۲-۳-۵- الگوی رگرسیونی خطی چند متغیره با متغیر های توضیحی تصادفی

در اینجا بر خلاف رگرسیون چند متغیره که در آن متغیر مستقل ثابت فرض شده است ، متغیر های توضیحی نیز تحت کنترل محقق نیستند،لذا تصادفی می باشند و ما در هر مرحله ای که می خواهیم آزمون فرض و تحلیل های دیگر را انجام دهیم، بردار $(y_1, y_2, \ldots, y_p, x_1, x_2, \ldots, x_q)$ را در اختیار داریم که همگی متغیر تصادفی اند.البته ما در این پروژه به بررسی این حالت نمی پردازیم و تنها توجه داشته باشد که تمامی محاسبات آزمون فرضها و برآورد ها مانند متغیر های ثابت انجام میشود و ما آن را برای مطالعه بیشتر به کتاب Rencher (1998, Section 7.7) ارجاع می دهیم .

۲-۴- آزمون فرض ها

هدف ما در این بخش تعمیم آزمون فرض های مربوط به رگرسیون خطی چندگانه به رگرسیون چند متغیره باشد.

۲-۴-۱- آزمون های رگرسیون کلی

در ابتدا ما می خواهیم فرض اینکه هیچ یک از x ها هیچ یک از y ها را پیش بینی نمی کند ، یا بعبارتی می خواهیم $H_0 : \mathbf{B}_1 = \mathbf{O}$ در مقابل $H_1 : \mathbf{B}_1 \neq \mathbf{O}$ را آزمون کنیم.

$$\begin{cases} H_0 : B_1 = 0 \\ H_1 : B_1 \neq 0 \end{cases}$$

$$\mathbf{B} = \begin{pmatrix} \boldsymbol{\beta}_0' \\ \mathbf{B}_1 \end{pmatrix} = \left(\begin{array}{cccc} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \hline \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \vdots & \vdots & & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{array} \right)$$

رابطه (۲۴)

ماتریس $\mathbf{Y}'\mathbf{Y}$ را می توان بصورت زیر افراز می کنیم:

$$\mathbf{Y}'\mathbf{Y} = (\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}) + \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}$$

رابطه (۲۵)

طبق توضیحاتی که در منبع ۱ صفحه ۳۳۰ آمده است می توان از دو طـرف رابطـه ۲۵ مقـدار $n\overline{\mathbf{y}}\,\overline{\mathbf{y}}'$ را کم می کنیم و در این صورت داریم :

رابطه (۲۶)

$$\begin{aligned} \mathbf{Y}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}' &= (\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}) + (\hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}') \\ &= \mathbf{E} + \mathbf{H}. \end{aligned}$$

حال در رگرسـیون کلـی بـا اسـتفاده از مـاتریس هـای E,H مـی تـوان آزمـون $H_0:\ \mathbf{B}_1 = \mathbf{O}$ را انجام داد[7].

بنابراین آماره آزمون مربوطه بصورت زیر بدست می آید :

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E} + \mathbf{H}|} = \frac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}'|},$$

رابطه (۲۷)

و تحت فرض $H_0:\ \mathbf{B}_1 = \mathbf{O}$ ، Λ دارای توزیع $\Lambda_{p,q,n-q-1}$ می باشد.

[7] برای مطالعه بیشتر به همان منبع فصل ۶ مراجعه کنید .

به آماره بالا اصطلاحاً لاندای ویلکس می گویند .فرض $H_0: \mathbf{B}_1 = \mathbf{O}$ رد مـی شـود اگـر $\Lambda \leq \Lambda_{\alpha,p,q,n-q-1}$ برقـرار باشـد .در صـورتی کـه از آزمـون هـای نـسبت راستنمایی استفاده کنیم ،می توان نشان داد که به چنین آماره ای می رسیم .

توجه کنید زمانی که مقدار $\hat{\beta}_{jk}$ ها بـزرگ هـستند مقـدارH بـزرگ مـی شـود،آنگاه مقـدار $|\mathbf{E} + \mathbf{H}|$ نسبت به $|\mathbf{E}|$ بزرگ شده و بنابراین مقدار Λ کوچک می شود کـه باعـث رد فرض $H_0: \mathbf{B}_1 = \mathbf{O}$ می شود. مقادیر بحرانی Λ در جدول a پیوست آمده اسـت کـه در آن مقدار $\nu_H = q$ و همچنین $\nu_E = n - q - 1$ می باشد .

توجـه کنیـد کـه درجـه آزادی ماننـد آزمـون یـک متغیـره رگرسـیون y بـر روی $x_1, x_2, \ldots, x_q$ می باشد . می توان تقریب F و χ^2 را نیز برای آماره لانـدای ویلکـس استفاده کرد(پیوست A) .

دو شـکل متفـاوت دیگـر بـرای محاسـبه Λ وجـود دارد ، کـه در آن از مقـادیر ویژه $\lambda_1, \lambda_2, \ldots, \lambda_s$ ماتریس $\mathbf{E}^{-1}\mathbf{H}$ استفاده میشود:

$$\Lambda = \prod_{i=1}^{s} \frac{1}{1 + \lambda_i},$$

که در آن $s = \min(p, q)$ می باشد .

و دیگری اگر S را بصورت زیر افراز کنیم :

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{yy} & \mathbf{S}_{yx} \\ \mathbf{S}_{xy} & \mathbf{S}_{xx} \end{pmatrix}$$

رابطه (۲۸)

آنگاه Λ بصورت :

$$\Lambda = \frac{|\mathbf{S}|}{|\mathbf{S}_{xx}||\mathbf{S}_{yy}|},$$

رابطه (۲۹)

در فرم اخیر که برای محاسبه Λ گفته شد شباهت زیادی با آماره ای که برای بررسی مستقل بودن دو بردار X و Y استفاده شد [8]،دارد. با این تفاوت که در آنجا هر دو بردار تصادفی بودند اما در اینجا ما فرض کرده ایم که بردار X ثابت است .بنابراین همانطور که در حالت عمومی داشتیم $\mathbf{S}_{yy}$ مانند یک ماتریس واریانس کوواریانس نمونه ای عمل می کند،و در حقیقت S_{xx} یک ماتریس شامل عبارات ریاضی قابل مقایسه هستند که دارای ثابت های x می باشد[9].

بدلیل تقارن x و y می توان گفت توزیع $\Lambda_{q,p,n-p-1}$ مانند $\Lambda_{p,q,n-q-1}$ می باشد.بنابراین در انجام آزمون های مربوط به بردارهای β متفاوت ،باز هم مقدار آماره آزمون ثابت می ماند .

[8] ۱۰,۵۷ منبع ۲ را ببینید.

[9] توضیح مربوط به S_{xx} را در بخش ۱۰,۲,۴ منبع ۲ ببینید.

مثال ۲-۲) فرض کنید برای داده های مثال ۲-۱ بخواهیم آزمون رگرسیون کلی $H_0: \mathbf{B}_1 = \mathbf{O}$ را انجام دهیم.در این صورت مقادیر ماتریس های H و E بصورت زیر بدست می آیند :

$$\mathbf{E} = \begin{pmatrix} 80.174 & -21.704 & -65.923 \\ -21.704 & 249.462 & -179.496 \\ -65.923 & -179.496 & 231.197 \end{pmatrix},$$

$$\mathbf{H} = \begin{pmatrix} 1707.158 & -492.532 & -996.584 \\ -492.532 & 151.002 & 283.607 \\ -996.584 & 283.607 & 583.688 \end{pmatrix}.$$

مقادیر ویژه ماتریس $\mathbf{E}^{-1}\mathbf{H}$ عبارتند از ۲۶,۳۱۸۳ ، ۰,۱۰۰۴ و ۰,۰۰۳۳ می باشد و مقدار آماره بصورت زیر بدست می آید:

$$\nu_H = q = 3, \qquad \nu_E = n - q - 1 = 19 - 3 - 1 = 15,$$

$$s = \min(3, 3) = 3,$$

$$\Lambda = \prod_{i=1}^{3} \frac{1}{1+\lambda_i} = \frac{1}{1+26.3183}\,\frac{1}{1+.1004}\,\frac{1}{1+.0033}$$
$$= .0332 < \Lambda_{.05,3,3,15} = .309,$$

۲-۴-۲- آزمون های مربوط به مجموعه ای از x ها :

در این قسمت می خواهیم آزمون کنیم که آیا y با h متغیر آخر یعنی x_{q-h+1}، $x_{q-h+2}, \ldots, x_q$ همبستگی دارد یا خیر ؟

برای این منظور ماتریس B را بصورت زیر افزار کرده و طریقه انجام آزمون $H_0: \mathbf{B}_d = \mathbf{O}.$ را توضیح می دهیم .

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_r \\ \mathbf{B}_d \end{pmatrix}$$ رابطه (۳۰)

$$X = [X_r \mid X_d]$$

و این نمایانگر این است که تعداد r متغیر در مدل کاهش یافته (Reduce model) می باشد و تعداد d متغیر از مدل کامل (Full model) حذف می شود (Deleted). می خواهیم آزمون کنیم که آیا این متغیر های حذف شده در پیش بینی y موثرند یا خیر ؟

بنابراین فرض $H_0: \mathbf{B}_d = \mathbf{O}.$ در صورت پذیرفته شدن ، مدل را بصورت $\mathbf{Y} = \mathbf{X}_r\mathbf{B}_r + \Xi$ تایید می کند.برای مقایسه مدل برازش داده شده و مدل کامل ما احتیاج به تفاوت مجموع مربعات رگرسیونی مدل کاهش یافته و مدل کامل داریم که بصورت زیر بدست می آید :

$$\mathbf{H} = \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}.$$ رابطه (۳۱)

برای ساختن آماره آزمون، ماتریس E بصورت زیر از مدل کامل بدست می آید :

$$\mathbf{E} = \mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}.$$ رابطه (۳۲)

پس

$$\begin{aligned}\mathbf{E} + \mathbf{H} &= (\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}) + (\hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}) \\ &= \mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y},\end{aligned}$$ رابطه (۳۳)

در اینصورت آماره لاندای ویلکس شرطیعبارت است از :

$$\begin{aligned}\Lambda(x_{q-h+1}, \ldots, x_q | x_1, \ldots, x_{q-h}) &= \frac{|\mathbf{E}|}{|\mathbf{E} + \mathbf{H}|} \\ &= \frac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}|},\end{aligned}$$

و زمانی که فرض $H_0: \mathbf{B}_d = \mathbf{O}.$ درست باشد دارای توزیع $\Lambda_{p,h,n-q-1}$ می باشد که مقادیر بحرانی آن در جدول a پیوست آمده است. توجه کنید که مثل قبل درجه آزادی مانند آزمون یک متغیره رگرسیون y بر روی $x_1, x_2, \ldots, x_q$ می باشد . می توان تقریب F و χ^2 را نیز برای آماره لاندای ویلکس استفاده کرد .

همانطور که در نماد گذاری بالا گفته آماره لاندای ویلکس بصورت $\Lambda(x_{q-h+1}, \ldots, x_q | x_1, \ldots, x_{q-h})$ بوده و بر اساس هر دو مدل کاهش یافته و کامل بیان می شود. و می توان مدل کامل را بر حسب مدل کاهش یافته بیان کرد.در نماد گذاری های قبل مقدار $\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}$ بیانگر مقدار مجموع مربعات خطا در مدل

کاهش یافته $Y = X_r B_r + \Xi$ است .همچنین از این مقدار خطا می توان برای آزمون فـرض رگرسیون کلی در مدل کاهش یافته نیز استفاده کرد .

$$\Lambda_r = \frac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}'|}$$

رابطه (۳۴)

که در آن همان نماد گذاری وجود دارد که در رابطـه ۲۷ آمـده بـود.بنـابراین آمـاره لانـدای ویلکس شرطی در آزمون بالا را می توان بصورت نسبت لانداهای ویلکـس در مـدل کامـل و مدل کاهش یافته (رگرسیون کلی) تبدیل کرد:

$$\begin{aligned}\Lambda(x_{q-h+1},\ldots,x_q|x_1,\ldots,x_{q-h}) &= \frac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}|}\\ &= \frac{\dfrac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}'|}}{\dfrac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_r'\mathbf{X}_r'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}'|}}\\ &= \frac{\Lambda_f}{\Lambda_r},\end{aligned}$$

رابطه (۳۵)

که Λ_f همان رابطه ۲۷ است و Λ_r رابطه ۳۴ می باشد. پس با توجه بـه توضـیحات بـالا مـا آزمون رگرسیون کلی را برای مدل کامل و همچنین مدل کاهش یافته بدسـت آورده و نـسبت آنها را بعنوان لاندای ویلکس شرطی در نظر می گیریم.

۲-۵- اندازه گیری رابطه بین x ها و y ها

ما درمبحث چند متغیره پیوسته با توزیع نرمال (در درس چند متغیره پیوسته) مقدار R^2 را بصورت $R^2 = \mathbf{s}'_{yx}\mathbf{S}_{xx}^{-1}\mathbf{s}_{yx}/s_{yy}$ محاسبه کردیم؛اما در اینجا چون بردار y تبدیل به ماتریس Y شده است داریم:

$$R_M^2 = \frac{|\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}|}{|\mathbf{S}_{yy}|}, \qquad \text{رابطه (۳۶)}$$

که در آن مقادیر $\mathbf{S}_{yx}, \mathbf{S}_{xy}, \mathbf{S}_{xx}$ و $\mathbf{S}_{yy}$ در رابطه ۲۹ تعریف شده است و نماد M در آن برای چند متغیره می باشد.

لازم به ذکر است همانطور که در قسمت روش های چند متغیره پیوسته خوانده شده است در اینجا نیز چند معیار همبستگی برای این منظور بوسیله افراد مختلفی پیشنهاد شده است ،که ما از ذکر همه آنها خودداری می کنیم[10].باید توجه داشت در اینجا نیز مانند قبل R^2 بین ۰ و ۱ تغییر می کند .

در زیر دو معیار همبستگی که بر حسب لاندای ویلکس بدست می آید، آمده است :

$$\eta_\Lambda^2 = 1 - \Lambda,$$
$$A_\Lambda = 1 - \Lambda^{1/s}, \qquad \text{رابطه (۳۷)}$$

که در آن $s = \min(p, q)$ می باشد .

[10] برای اطلات بیشتر به منبع ۲ فصل ۶ مراجعه کنید

مثال ۲-۳) از داده های مثال ۱ (جدول ۱) برای بررسی مفهوم بالا و محاسبه چنـدین کمیـت رابطه ای توضیح داده شده در بالا،استفاده می کنیم. در این مثال سه متغیر وابسته و ســه متغیــر مستقل وجود دارد و ماتریس واریانس کوواریانس بصورت زیر افراز می شود :

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{yy} & \mathbf{S}_{yx} \\ \mathbf{S}_{xy} & \mathbf{S}_{xx} \end{pmatrix}$$

$$= \left(\begin{array}{rrr|rrr} 99.30 & -28.57 & -59.03 & -41.95 & -9.49 & -7.37 \\ -28.57 & 22.25 & 5.78 & 11.85 & 1.60 & 3.03 \\ -59.03 & 5.78 & 45.27 & 24.14 & 6.43 & 3.97 \\ \hline -41.95 & 11.85 & 24.14 & 38.67 & -12.17 & -.22 \\ -9.49 & 1.60 & 6.43 & -12.17 & 17.95 & 1.22 \\ -7.36 & 3.03 & 3.97 & -.22 & 1.22 & 2.67 \end{array}\right)$$

و بنابراین مقادیر زیر بدست می آید :

$$R_M^2 = \frac{|\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}|}{|\mathbf{S}_{yy}|} = .00029$$

$$\eta_\Lambda^2 = 1 - \Lambda = 1 - .0332 = .967,$$

$$A_\Lambda = 1 - \Lambda^{1/s} = 1 - \Lambda^{1/3} = .679$$

۲-۶- انتخاب مجموعه ها

در اینجا نیز مانند حالت یک متغیره (رگرسیون خطی چندگانه)که ممکن بود چندین x داشته باشیم که تاثیری در پیش بینی y ها نداشته باشند، در رگرسیون چند متغیره نیز ممکن است بعضی از x ها تاثیری در پیش بینی برخی از y ها نداشته باشند و لازم باشد آن ها را از مدل حذف کنیم.در اینجا ما دو حالت را برای انتخاب مجموعه ها بررسی می کنیم، یکی روش گام به گام و دیگری روش کلی که تمام مجموعه های ممکن را بررسی می کند ،که ما در این پروژه به روش گام به گام اکتفا میکنیم .

۲-۶-۱- روش گام به گام (Stepwise)

روش انتخاب گام به گام برای x ها ؛که شامل دو روش پسرو[۱۱] و پیشرو[۱۲] می باشد؛ در بخش ۲-۶-۱-۱ توضیح داده می شود و در بخش ۲-۶-۱-۲ این روش را برای y مورد بررسی قرار می دهیم.

۲-۶-۱-۱- یافتن مجموعه ای از x ها

در ابتدا با روش پیشرو و با استفاده از آماره ی لاندای ویلکس اثر تک به تک x ها را آزمون کرده،بنابراین ماتریس B بصورت دو سطر در می آید که سطر اول شامل پارامتر های ضریب ثابت و سطر دوم پارامترهای ضریب متغیر jام (یا بعبارتی β_{ji} ها) است. پس داریم :

[11] backward

[12] forward

$$\hat{\mathbf{B}}_j = \begin{pmatrix} \hat{\beta}_{01} & \hat{\beta}_{02} & \cdots & \hat{\beta}_{0p} \\ \hat{\beta}_{j1} & \hat{\beta}_{j2} & \cdots & \hat{\beta}_{jp} \end{pmatrix}$$
رابطه (۳۸)

با استفاده از آزمون رگرسیون کلی مقادیر $\Lambda(x_j)$ را حساب کرده و متغیری که مقـدار مینـیمم $\Lambda(x_j)$ را داشته باشد ، وارد مدل می شود.

در زیر آماره $\Lambda(x_j)$ و توزیع آن آمده است.

$$\Lambda(x_j) = \frac{|\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}_j'\mathbf{X}_j'\mathbf{Y}|}{|\mathbf{Y}'\mathbf{Y} - n\overline{\mathbf{y}}\,\overline{\mathbf{y}}'|}$$
رابطه (۳۹)

که دارای توزیع $\Lambda_{p,1,n-2}$ می باشد زمانی کـه B دارای دو سـطر و x_j دارای دو سـتون مـی باشد.

در قدم دوم ما متغیری را که در حضور متغیر اولی که وارد شده است، کمترین لاندای شـرطی را داراست وارد مدل می کنیم،که در آن لاندای شرطی ویلکس بصورت زیر تعریف می شود :

$$\Lambda(x_j|x_1) = \frac{\Lambda(x_1, x_j)}{\Lambda(x_1)}$$
رابطه (۴۰)

که در آن x_1 نمایانگر اولین متغیری است که وارد مدل شده است .

و بهمین ترتیب سومین متغیر با بررسی لاندای شرطی ویلکس (البته بـه شـرط دو متغیـر وارد شده) متغیری که کمترین لاندا را داراست وارد مدل می شود .

$$\Lambda(x_j|x_1, x_2) = \frac{\Lambda(x_1, x_2, x_j)}{\Lambda(x_1, x_2)}.$$
رابطه (۴۱)

و بهمین ترتیب m امین متغیر را با بررسی لاندای زیر و مینیمم کردن آن وارد مدل می کنیم :

$$\Lambda(x_j|x_1, x_2, \ldots, x_m) = \frac{\Lambda(x_1, x_2, \ldots, x_m, x_j)}{\Lambda(x_1, x_2, \ldots, x_m)}$$

رابطه (۴۲)

این فرآیند به همین شکل ادامه پیدا می کند،و قدم به قدم بهترین متغیر ها کــه دارای کمتــرین لاندای ویلکس و یا بعبارتی بیشترین وابستگی را برای پیش بینی y ها دارند؛را وارد مــدل مــی کند.

برای متوقف شدن این فرآیند بایستی در هر حالت مقدار-p تعیین شده را با مقــدار-p لانــدای ویلکس مقایسه کرده و در صورتی که مقدار-p لاندای ویلکس (مینــیمم لانــداها)از مقــدار-p تعیین شده بیشتر باشد ، فرآیند متوقف می شود. توجه کنید که در هر حالت که متغیــر x وارد مدل می شود احتمال دارد که یک سری تاثیر های متقابل بین آنها بوجود آید که در این حالت لازم می شود بعضی از آنها را از مدل حذف کرد.

در روش گام به گام که تنها تعمیمی از روش پیشرو است،در هر مرحلــه کــه یــک متغیر وارد مدل میشود، به بررسی تک تک متغیر های وارد مدل شده می پردازیم و با توجه بــه سطح معنی داری تعیین شده از ویلکس جزئی؛آن متغیری را که فرض صفر بــودن آن پذیرفتــه می شود از مدل حذف می کنیم(یا بعبارتی مقدار سطح معنی داری آن بیــشتر از ســطح معنــی داری تعیین شده باشد).

در روش پسرو ابتدا همه متغیر های x وارد مدل می شوند(همه سطر های B) و در ابتدا با استفاده آماره لاندای ویلکس جزئی که در زیر آمده است، آن متغیـری را کـه بیـشترین لاندای ویلکس را داراست از مدل حذف می شود.

$$\Lambda(x_j|x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_q) = \frac{\Lambda(x_1, \ldots, x_q)}{\Lambda(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_q)},$$

رابطه (۴۳)

که دارای توزیع $\Lambda_{p,1,n-q-1}$ می باشد .

و در قدم دوم نیز به همین ترتیب q-1 متغیر باقیمانده را بررسی کرده و آن متغیـری را که بیشترین لاندا را داراست از مدل حذف می شود و ... .

زمانی این فرآیند متوقف می شود که مقدار سطح معنی داری لاندای ویلکس جزئی ماکزیمم ،کمتر از سطح معنی داری تعیین شده باشد.ممکن است مقداری از آماره لاندا نیز بـه عنوان شرط توقف تعیین شود.

توجه داشته باشد که ممکن است در روش گام به گام مدل بهینه را بدست نیـاوریم و حتی ممکن است مقدار بهینه لاندای ویلکس که از روش بدست آمده با مقدار واقعی تفاوت زیادی کند،لذا بایستی روش های متفاوت را بررسی کرده و خودمان در بین آنها بهتـرین مـدل را پیدا کنیم (که البته نیاز به تجربه و انجام کارهای عملی بسیاری دارد).

۲-۶-۱-۲- یافتن مجموعه ای از y ها

بعد از یافتن مجموعه بهینه ای از x ها ، ممکن است محقق خواستار این باشد که آیا p متغیر y توسط x ها پیش بینی می شوند یا خیر.اگر تعدادی از y ها هیچ رابطه ای با هیچ کدام از متغیر های مستقل نداشته باشند،بایستی آنها را از مدل حذف کرد.

طبق توضیحات گفته در مورد انتخاب مجموعه از x ها در بخش قبل میتوان از آن روش برای این سری متغیر ها نیز استفاده کرد (رابطه های ۴۰،۴۱،۴۲و۴۳) .

لذا در اینجا نیز بعنوان مثال در روش پیشرو اگر دو متغیر وارد شده y_1 وy_2 باشند ، لاندای ویلکس شرطی بصورت :

$$\Lambda(y_j|y_1, y_2) = \frac{\Lambda(y_1, y_2, y_j)}{\Lambda(y_1, y_2)}$$

رابطه (۴۴)

لاندایی که کمترین مقدار را دارد متغیر مربوط به آن وارد مدل می شود.

همچنین در اولین قدم روش پسرو مقدار لاندای زیر را محاسبه و مقدار ماکزیمم آنها را پیدا کرده و از مدل متغیر مربوط به آن را حذف می کنیم.

$$\Lambda(y_j|y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_p) = \frac{\Lambda(y_1, \ldots, y_p)}{\Lambda(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_p)}$$

رابطه (۴۵)

که دارای توزیع $\Lambda_{1,q,n-q-p}$ می باشد.و در این روش ها نیز مانند x شرط توقف می تواند بسته به شرط معنی داری و یا مقدار از پیش تعیین شده باشد.

فصل سوم

تحلیل های آماری

با نرم افزار SAS

فصل سوم : تحلیل های آماری با نرم افزار SAS

۳–۱– هدف

در این فصل می خواهیم با استفاده از نرم افزار SAS ، خروجی برنامه و تحلیل های آماری مربوط به رگرسیون چند متغیره را توضیح دهیم.برای این منظور با استفاده از منبع ۳ مثال های مربوط به قسمت های مختلف را آورده و با تغییرات جزئی روی برنامه ها تحلیل های مربوط به خروجی ها را توضیح می دهیم.

۳–۲– محاسبات مربوط به آزمون فرض ها

مثال ۳–۱–

برای یادگیری بیشتر در آزمون فرضها مثال مربوط به کتاب Carter (1983) با نام داده های ماهی را در اینجا آورده و تحلیل های لازم را با استفاده از برنامه ۳–۱ انجام می دهیم. در این مثال می خواهیم تاثیر فلز مس را بر مرگ و میر ماهی ها بررسی کنیم.برای این منظور در ۲۵ حوض ماهی که هر کدام شامل ۲۰ ماهی قزل آلا بوده است غذاهای مختلفی که وزن های مس آنها(میلی گرم در لیتر) متفاوت بوده است،داده ایم و نسبت ماهی های مرده در پنج زمان متفاوت۸ ، ۱۴ ، ۲۴، ۳۶ و ۴۸ (مقادیر P1 تا P5) را اندازه گرفته ایم. برای پایدار سازی واریانس مقدار arcsin جذر آنها را بعنوان Y1 تا Y5 در نظر می گیریم.همچنین وزن ماهی های هر حوض و مقدار مس موجود درخوراک این ۲۵ حوض (DOSE) ؛بر حسب

میلی گرم در لیتر؛ بعنوان متغیر های مستقل آورده شده است. بعلاوه برای حفظ شیوه استاندارد از log(DOSE) بعنوان متغیر مستقل استفاده می کنیم،این چنین تبدیل هایی تغییرات بین متغیر ها را بصورت یکنواخت تبدیل می کند.

اهداف ما در این مسئله عبارتند از :

- برازش دادن یک مدل چند متغیره روشن برای Yi ها با استفاده از تابعی خطی وزن ها و لگاریتم طبیعی dose
- آزمون سطح معنی داری برای تاثیر dose
- آزمون سطح معنی داری برای تاثیر وزن(Weight)
- آزمون معنی داری برای کل مدل

P1	P2	P3	P4	P5	WHEITH	DOSE
0	0	.25	.25	.25	270	.6695
0	.10	.30	.30	.30	410	.6405
0	.5	.75	.9	.9	610	.729
.15	.65	10	10	10	940	.77
.45	1	1	1	1	1450	.5655
0	.05	.20	.20	.2	270	.782
.05	.1	.3	.3	.3	410	.812
.05	.45	.95	1	1	610	.8215
.1	.7	1	1	1	940	.869
.2	.85	1	1	1	1450	.8395
0	0		0	.05	270	.8615
0	.05	.15	.25	.30	410	.9045
0	.15	.95	.95	.95	610	1028
0	.55	.95	1	1	940	10445
.1	.85	1	1	1	1450	10455
0	0	0	.05	.10	270	.6195
0	.05	.15	.20	.25	410	.5305
.1	.45	.95	.95	.95	610	.597
.1	.7	1	1	1	940	.6385
.35	.95	1	1	1	1450	.6645
0	.05	.20	.20	.20	270	.5685
0	0	.15	.25	.25	410	.604
0	.4	.9	1	1	610	.6325
.05	.65	1	1	1	940	.6845
.3	.85	1	1	1	1450	.723

جدول ۲

حال می خواهیم با استفاده از نرم افزار SAS آزمون های مذکور را انجام دهیم.برای این منظور با استفاده از برنامه ۳-۱ مقادیر Y1 الی Y5 را از تبدیل Yi=arcsin[sqrt(pi)] بدست می آوریم.وهمچنین متغیر x1= log (dose) و متغیر X2 را برابر وزن ماهی ها در نظر می گیریم.

در برنامه ۳-۱ می خواهیم مدل زیر را برازش دهیم:

$$\mathbf{Y}_{25\times5} = \mathbf{X}_{25\times3}\mathbf{B}_{3\times5} + \mathcal{E}_{25\times5},$$ رابطه (۴۶)

که در آن

$$\mathbf{B} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

و توجه کنید که مقادیر ۱ تا ۵ سطر بردار β_0 مقادیر ثابت برای مدل می باشند و همچنین مقادیر β_1,β_2 بعنوان پارامتر های شیب مدل برای متغیر های مستقل X1 و X2 می باشند.

برای آزمون سطح معنی داری در ابتدا فرض رگرسیون کلی زیر :

$$H_0 : \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \mathbf{0}$$

را بررسی می کنیم.که در اینجا فرض کرده ایم که خطاهای مدل دارای توزیع نرمال چند متغیره هستند،و با استفاده از لاندای ویلکس آزمون را انجام می دهیم.

برای بدست آوردن این آماره آزمون نمی توان از تابع PROC GLM استفاده کرد، هرچند بعدا ما با استفاده از برنامه PROC REG این محاسبات را توضیح می دهیم.

در این حالت با استفاده از

```
manova h=x1 x2;
```

مـی تـوانیم بـرای آزمـون فـرض هـای $H_0^{(1)} : \beta_1 = \mathbf{0}$ (عـدم تـاثیر متغیـر dose) و $H_0^{(2)} : \beta_2 = \mathbf{0}$ (عدم تاثیر وزن) محاسبات لازم را انجام دهیم.

با استفاده از تنظیمات PRINTE & PRINTH مقادیر مجموع مربعـات و مـاتریس بـرآورد خطاها را بدست می آوریم.

ابتدا با استفاده از دستور زیر داده ها را وارد می کنیم.

```
data work.fish;
 input p1 p2 p3 p4 p5 dose wt ;
 y1=arsin(sqrt(p1));
 y2=arsin(sqrt(p2));
 y3=arsin(sqrt(p3));
 y4=arsin(sqrt(p4));
 y5=arsin(sqrt(p5));
 x1=log(dose);
 x2=wt;
 cards;
 0 0. .25 .25 .25 270 .6695
 0. .10 .30 .30 .30 410 .6405
 0. .5 .75 .9 .9 610 .729
 .15 .65 1.0 1.0 1.0 940 .77
 .45 1. 1. 1. 1. 1450 .5655
 0. .05 .20 .20 .2 270 .782
 .05 .1 .3 .3 .3 410 .812
 .05 .45 .95 1. 1. 610 .8215
 .1 .7 1. 1. 1. 940 .869
 .2 .85 1. 1. 1. 1450 .8395
 0. 0. 0. 0. .05 270 .8615
 0. .05 .15 .25 .30 410 .9045
 0. .15 .95 .95 .95 610 1.028
 0. .55 .95 1. 1. 940 1.0445
 .1 .85 1. 1. 1. 1450 1.0455
 0. 0. 0. .05 .10 270 .6195
 0. .05 .15 .20 .25 410 .5305
 .1 .45 .95 .95 .95 610 .597
 .1 .7 1 1 1 940 .6385
 .35 .95 1. 1. 1. 1450 .6645
 0 .05 .20 .20 .20 270 .5685
 0 0 .15 .25 .25 410 .604
 0 .4 .9 1. 1. 610 .6325
 .05 .65 1 1. 1. 940 .6845
 .3 .85 1. 1. 1. 1450 .723
 ;
```

برنامه ۳–۱–

```
options ls=64 ps=45 nodate nonumber;
title1 'Output 3.1';

proc print data=fish;
var p1 p2 p3 p4 p5 dose x2;
title2 'Data on Proportions of Dead Fish';
run;
proc print data=fish;
var y1 y2 y3 y4 y5 x1 x2;
title2 'Transformed Fish Data';
run;
proc glm data=fish;
model y1 y2 y3 y4 y5=x1 x2/nouni;
manova h=x1 x2/printe printh;
title2 'Multivariate Regression for Fish Data';
run;
/*
mtest option of proc reg can be used instead of manova
option of proc glm to get the same results. This is done
using the last two statements of the following program.
*/
proc reg data=fish;
model y1 y2 y3 y4 y5=x1 x2;
Model: mtest x1, x2/print;
Onlyx1: mtest x1/print;
Onlyx2: mtest x2/print;
```

خروجی برنامه ۳-۱-

پس با استفاده از برنامه ۳-۱ محاسبات لازم را انجام داده و خروجی های آن بصورت زیر مـی باشند :

'Output 3.1'
'Data on Proportions of Dead Fish'

Obs	p1	p2	p3	p4	p5	dose	x2
1	0.00	0.00	0.25	0.25	0.25	270	0.6695
2	0.00	0.10	0.30	0.30	0.30	410	0.6405
3	0.00	0.50	0.75	0.90	0.90	610	0.7290
4	0.15	0.65	1.00	1.00	1.00	940	0.7700
5	0.45	1.00	1.00	1.00	1.00	1450	0.5655
6	0.00	0.05	0.20	0.20	0.20	270	0.7820
7	0.05	0.10	0.30	0.30	0.30	410	0.8120
8	0.05	0.45	0.95	1.00	1.00	610	0.8215
9	0.10	0.70	1.00	1.00	1.00	940	0.8690
10	0.20	0.85	1.00	1.00	1.00	1450	0.8395
11	0.00	0.00	0.00	0.00	0.05	270	0.8615
12	0.00	0.05	0.15	0.25	0.30	410	0.9045
13	0.00	0.15	0.95	0.95	0.95	610	1.0280
14	0.00	0.55	0.95	1.00	1.00	940	1.0445
15	0.10	0.85	1.00	1.00	1.00	1450	1.0455
16	0.00	0.00	0.00	0.05	0.10	270	0.6195
17	0.00	0.05	0.15	0.20	0.25	410	0.5305
18	0.10	0.45	0.95	0.95	0.95	610	0.5970
19	0.10	0.70	1.00	1.00	1.00	940	0.6385
20	0.35	0.95	1.00	1.00	1.00	1450	0.6645
21	0.00	0.05	0.20	0.20	0.20	270	0.5685
22	0.00	0.00	0.15	0.25	0.25	410	0.6040
23	0.00	0.40	0.90	1.00	1.00	610	0.6325
24	0.05	0.65	1.00	1.00	1.00	940	0.6845
25	0.30	0.85	1.00	1.00	1.00	1450	0.7230

'Output 3.1'
'Transformed Fish Data'

Obs	y1	y2	y3	y4	y5	x1	x2
1	0.00000	0.00000	0.52360	0.52360	0.52360	5.59842	0.6695
2	0.00000	0.32175	0.57964	0.57964	0.57964	6.01616	0.6405
3	0.00000	0.78540	1.04720	1.24905	1.24905	6.41346	0.7290
4	0.39770	0.93774	1.57080	1.57080	1.57080	6.84588	0.7700
5	0.73531	1.57080	1.57080	1.57080	1.57080	7.27932	0.5655
6	0.00000	0.22551	0.46365	0.46365	0.46365	5.59842	0.7820
7	0.22551	0.32175	0.57964	0.57964	0.57964	6.01616	0.8120
8	0.22551	0.73531	1.34528	1.57080	1.57080	6.41346	0.8215
9	0.32175	0.99116	1.57080	1.57080	1.57080	6.84588	0.8690
10	0.46365	1.17310	1.57080	1.57080	1.57080	7.27932	0.8395
11	0.00000	0.00000	0.00000	0.00000	0.22551	5.59842	0.8615
12	0.00000	0.22551	0.39770	0.52360	0.57964	6.01616	0.9045
13	0.00000	0.39770	1.34528	1.34528	1.34528	6.41346	1.0280
14	0.00000	0.83548	1.34528	1.57080	1.57080	6.84588	1.0445
15	0.32175	1.17310	1.57080	1.57080	1.57080	7.27932	1.0455
16	0.00000	0.00000	0.00000	0.22551	0.32175	5.59842	0.6195
17	0.00000	0.22551	0.39770	0.46365	0.52360	6.01616	0.5305
18	0.32175	0.73531	1.34528	1.34528	1.34528	6.41346	0.5970
19	0.32175	0.99116	1.57080	1.57080	1.57080	6.84588	0.6385
20	0.63305	1.34528	1.57080	1.57080	1.57080	7.27932	0.6645
21	0.00000	0.22551	0.46365	0.46365	0.46365	5.59842	0.5685
22	0.00000	0.00000	0.39770	0.52360	0.52360	6.01616	0.6040
23	0.00000	0.68472	1.24905	1.57080	1.57080	6.41346	0.6325
24	0.22551	0.93774	1.57080	1.57080	1.57080	6.84588	0.6845
25	0.57964	1.17310	1.57080	1.57080	1.57080	7.27932	0.7230

The GLM Procedure

Number of Observations Read	25
Number of Observations Used	25

'Multivariate Regression for Fish Data'

The GLM Procedure
Multivariate Analysis of Variance

E = Error SSCP Matrix

	y1	y2	y3
y1	0.3521207721	0.1323723295	-0.081582595
y2	0.1323723295	0.3986006993	0.2479046457
y3	-0.081582595	0.2479046457	1.2273557756
y4	-0.238505417	0.2352986084	1.2221331781
y5	-0.211550264	0.2252678955	1.0891872222

E = Error SSCP Matrix

	y4	y5
y1	-0.238505417	-0.211550264
y2	0.2352986084	0.2252678955
y3	1.2221331781	1.0891872222
y4	1.4505775203	1.3142693993
y5	1.3142693993	1.2243378999

Partial Correlation Coefficients from
the Error SSCP Matrix / Prob > |r|

DF = 22	y1	y2	y3	y4	y5
y1	1.000000	0.353332 0.0982	-0.124098 0.5726	-0.333720 0.1197	-0.322194 0.1338
y2	0.353332 0.0982	1.000000	0.354430 0.0970	0.309442 0.1508	0.322463 0.1334
y3	-0.124098 0.5726	0.354430 0.0970	1.000000	0.915931 <.0001	0.888519 <.0001
y4	-0.333720 0.1197	0.309442 0.1508	0.915931 <.0001	1.000000	0.986196 <.0001
y5	-0.322194 0.1338	0.322463 0.1334	0.888519 <.0001	0.986196 <.0001	1.000000

'Multivariate Regression for Fish Data'

The GLM Procedure
Multivariate Analysis of Variance

H = Type III SSCP Matrix for x1

	y1	y2	y3
y1	0.9444331687	2.1396606906	2.4004527379
y2	2.1396606906	4.8475085618	5.4383460192
y3	2.4004527379	5.4383460192	6.1011975631
y4	2.3189700315	5.2537428628	5.8940940938
y5	2.2097384417	5.0062732199	5.6164616708

H = Type III SSCP Matrix for x1

	y4	y5
y1	2.3189700315	2.2097384417
y2	5.2537428628	5.0062732199
y3	5.8940940938	5.6164616708
y4	5.6940206946	5.4258124278
y5	5.4258124278	5.1702377073

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for x1
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 y1 / y3 / y5	y2 / y4
14.2472850	100.00	0.55171485	1.04785724
		0.06215532	-0.11654144
		0.46436590	
0.0000000	0.00	-0.31734319	-0.24878068
		0.72475369	-5.81185638
		5.68836531	

```
                         The GLM Procedure
                 Multivariate Analysis of Variance

   Characteristic Roots and Vectors of: E Inverse * H, where
                H = Type III SSCP Matrix for x1
                     E = Error SSCP Matrix

                                      Characteristic
   Characteristic                     Vector  V'EV=1
             Root      Percent                  y1             y2
                                                y3             y4
                                                y5

        0.0000000         0.00          1.31612722    -0.65966844
                                       -2.38942315     2.54603108
                                        0.00000000
        0.0000000         0.00          0.43144422    -1.25612329
                                        0.94990722     0.00000000
                                        0.00000000
        0.0000000         0.00         -1.78764520     0.78905568
                                        0.00000000     0.00000000
                                        0.00000000

          MANOVA Test Criteria and Exact F Statistics
          for the Hypothesis of No Overall x1 Effect
                H = Type III SSCP Matrix for x1
                     E = Error SSCP Matrix

                      S=1    M=1.5    N=8

Statistic                       Value    F Value    Num DF    Den DF

Wilks' Lambda              0.06558545      51.29         5        18
Pillai's Trace             0.93441455      51.29         5        18
Hotelling-Lawley Trace    14.24728504      51.29         5        18
Roy's Greatest Root       14.24728504      51.29         5        18
```

'Output 3.1'
'Multivariate Regression for Fish Data'

The GLM Procedure
Multivariate Analysis of Variance

MANOVA Test Criteria and Exact F Statistics
for the Hypothesis of No Overall x1 Effect
H = Type III SSCP Matrix for x1
E = Error SSCP Matrix

S=1 M=1.5 N=8

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

H = Type III SSCP Matrix for x2

	y1	y2	y3
y1	0.0976730369	0.0792504059	-0.01314736
y2	0.0792504059	0.0643025654	-0.010667567
y3	-0.01314736	-0.010667567	0.0017697113
y4	-0.020670718	-0.016771904	0.0027823991
y5	-0.032603961	-0.026454355	0.0043886832

H = Type III SSCP Matrix for x2

	y4	y5
y1	-0.020670718	-0.032603961
y2	-0.016771904	-0.026454355
y3	0.0027823991	0.0043886832
y4	0.0043745806	0.0069000341
y5	0.0069000341	0.0108834365

```
                         The GLM Procedure
                 Multivariate Analysis of Variance

  Characteristic Roots and Vectors of: E Inverse * H, where
                  H = Type III SSCP Matrix for x2
                       E = Error SSCP Matrix

                                    Characteristic
  Characteristic                    Vector  V'EV=1
            Root     Percent                    y1              y2
                                                y3              y4
                                                y5

      0.52693756      100.00            1.54528621      0.54470517
                                       -1.23989398      3.77134305
                                       -2.89593127
      0.00000000        0.00            0.70777693     -0.73863887
                                        0.90830216      0.30892057
                                       -0.23721330
      0.00000000        0.00            0.88973917     -0.71121690
                                       -2.13333142      1.69536789
                                        0.72207496
      0.00000000        0.00           -1.15517657      1.56318576
                                        0.08060993     -0.81221542
                                        0.82145754
      0.00000000        0.00            0.71005629     -0.06199478
                                        0.46994370     -4.73509285
                                        4.78897304


                            'Output 3.1'
                'Multivariate Regression for Fish Data'

                         The GLM Procedure
                 Multivariate Analysis of Variance

             MANOVA Test Criteria and Exact F Statistics
             for the Hypothesis of No Overall x2 Effect
                  H = Type III SSCP Matrix for x2
                       E = Error SSCP Matrix

                        S=1     M=1.5     N=8

Statistic                         Value    F Value    Num DF    Den DF

Wilks' Lambda                0.65490563       1.90         5        18
Pillai's Trace               0.34509437       1.90         5        18
Hotelling-Lawley Trace       0.52693756       1.90         5        18
Roy's Greatest Root          0.52693756       1.90         5        18

             MANOVA Test Criteria and Exact F Statistics
              for the Hypothesis of No Overall x2 Effect
                  H = Type III SSCP Matrix for x2
                       E = Error SSCP Matrix

                        S=1     M=1.5     N=8

                  Statistic                 Pr > F

                  Wilks' Lambda             0.1449
                  Pillai's Trace            0.1449
                  Hotelling-Lawley Trace    0.1449
                  Roy's Greatest Root       0.1449
```

همانطور که دیدید در خروجی ۳-۱ در ابتدا مجموع مربعات خطا و یا همان ماتریس E نمایش داده شده است.و مجموع مربعات بر حسب خطا بـرای هریـک از y_i هـا بـصورت برداری از این ماتریس می شود. پس از آن این خروجی تحلیل های مربوط بـه آزمـون فـرض x1 (لگاریتم طبیعی DOSE)را انجام می دهد.در این حالت بصورت پیش فـرض نـرم افـزار SAS حداقل مربعات خطای نوع سوم را برای آن محاسبه می کند و همچنین برای متغیر هـای دیگر نیز(در اینجا x2 ،وزن ماهی ها) این محاسبات انجام شده است.

همانطور که در خروجی ۳-۱ مشخص است همه ۴ آماره آزمـون بطـور مـشخص فرض $H_0^{(1)} : \beta_1 = 0$ (عدم تاثیر متغیر dose) را رد می کنند،یا بعبـارتی عامـل dose بطور شدیدی برای پیش بینی متغیر های وابسته لازم است.اما از طرف دیگر هر ۴ آماره آزمـون رابطه بین x2(وزن ماهی ها) و Yi ها را نقض می کنند،یا بعبارتی فـرض $H_0^{(2)} : \beta_2 = 0$ پذیرفته می شود.

در برنامه ۳-۱ در قسمت proc GLM نمی توان آزمون

$$H_0^{(3)} : \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

را انجام داد؛ در صورتی که یکی از آزمون های مهم در رگرسیون،آزمون رگرسیون کلـی مـی باشد که در آن آزمون می کند که آیا هر دو متغیر بطور همزمان وجود آنها در مـدل ضـروری است یا خیر؟ بنابراین با استفاده از برنامه proc REG و تابع MTEST می توان این آزمـون

را انجام داد .در این قسمت اگر بخـواهیم متغیـر ثابـت را در مـدل در نظـر گیـرد مـی تـوان INTERCEPT را بعد از MTEST قرار داد.

در برنامه ما این قسمت بصورت :

```
proc reg;
y1 y2 y3 y4 y5=x1 x2;
mtest x1, x2/print;
```

آورده شده است.که تنظیمات print برای این منظور است که ماتریس H و ماتریس خطاهای SS,CP (بعبارتی ماتریس E) را در خروجی چاپ کند.همانطور که در خروجـی ادامـه ۳-۱ مشاهده می کنید آماره لاندای ویلکس نیز محاسبه شده است و با استفاده از هر ۴ آماره آزمون می توان گفت که

$$H_0^{(3)} : \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \mathbf{0}$$

را می توان رد کرد.بنابراین می توان گفت نسبت ماهی های مرده با میانگین وزن مـاهی و یـا dose و یا هر دو وابستگی دارد.

همچنین با استفاده از MTEST می توان برای هر کدام از x ها نیز این آزمـون را انجام داد،بنابراین می توان در ایـن حالـت از MTEST در PROC REG همـانطور کـه از MANOVA در PROC GLM بکار بردیم ، استفاده کرد و خروجی هـای آن یکـسان مـی باشد.در ادامه خروجی ۳-۱ این مقادیر آمده است :

ادامه خروجی ۳–۱ (MTEST)

The REG Procedure
Model: MODEL1
Dependent Variable: y1

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	0.95388	0.47694	29.80
Error	22	0.35212	0.01601	
Corrected Total	24	1.30600		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.12651	R-Square	0.7304
Dependent Mean	0.19092	Adj R-Sq	0.7059
Coeff Var	66.26628		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-1.65110	0.28363	-5.82	<.0001
x1	1	0.33640	0.04379	7.68	<.0001
x2	1	-0.43079	0.17439	-2.47	0.0217

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Dependent Variable: y2

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	4.90810	2.45405	135.45
Error	22	0.39860	0.01812	
Corrected Total	24	5.30670		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.13460	R-Square	0.9249
Dependent Mean	0.64051	Adj R-Sq	0.9181
Coeff Var	21.01524		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-3.99987	0.30177	-13.25	<.0001
x1	1	0.76214	0.04659	16.36	<.0001
x2	1	-0.34954	0.18554	-1.88	0.0729

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Dependent Variable: y3

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	6.47479	3.23740	58.03
Error	22	1.22736	0.05579	
Corrected Total	24	7.70215		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.23620	R-Square	0.8406
Dependent Mean	1.02471	Adj R-Sq	0.8262
Coeff Var	23.05005		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-4.51696	0.52954	-8.53	<.0001
x1	1	0.85504	0.08176	10.46	<.0001
x2	1	0.05799	0.32558	0.18	0.8603

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Dependent Variable: y4

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	6.07427	3.03714	46.06
Error	22	1.45058	0.06594	
Corrected Total	24	7.52485		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.25678	R-Square	0.8072
Dependent Mean	1.08543	Adj R-Sq	0.7897
Coeff Var	23.65692		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-4.29435	0.57568	-7.46	<.0001
x1	1	0.82601	0.08889	9.29	<.0001
x2	1	0.09117	0.35395	0.26	0.7991

The REG Procedure
Model: MODEL1
Dependent Variable: y5

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	5.56711	2.78355	50.02
Error	22	1.22434	0.05565	
Corrected Total	24	6.79144		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.23591	R-Square	0.8197
Dependent Mean	1.10294	Adj R-Sq	0.8033
Coeff Var	21.38889		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-4.06589	0.52889	-7.69	<.0001
x1	1	0.78710	0.08166	9.64	<.0001
x2	1	0.14380	0.32518	0.44	0.6626

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Multivariate Test: Model

Error Matrix (E)

0.3521207721	0.1323723295	-0.081582595
0.1323723295	0.3986006993	0.2479046457
-0.081582595	0.2479046457	1.2273557756
-0.238505417	0.2352986084	1.2221331781
-0.211550264	0.2252678955	1.0891872222

Error Matrix (E)

-0.238505417	-0.211550264
0.2352986084	0.2252678955
1.2221331781	1.0891872222
1.4505775203	1.3142693993
1.3142693993	1.2243378999

Hypothesis Matrix (H)

0.9538786705	2.1157383555	2.3410489485
2.1157383555	4.9080959325	5.58879619
2.3410489485	5.58879619	6.4747944636
2.2590396342	5.4055267572	6.2710028881
2.1485124778	5.1613383564	6.0015184245

Hypothesis Matrix (H)

2.2590396342	2.1485124778
5.4055267572	5.1613383564
6.2710028881	6.0015184245
6.0742707424	5.8142826656
5.8142826656	5.567105838

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Multivariate Test: Model

Multivariate Statistics and F Approximations

S=2 M=1 N=8

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.04668792	13.06	10	36
Pillai's Trace	1.21319964	5.86	10	38
Hotelling-Lawley Trace	14.85233102	25.89	10	24.4
Roy's Greatest Root	14.46757522	54.98	5	19

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

NOTE: F Statistic for Wilks' Lambda is exact.

Multivariate Statistics and F Approximations

S=2 M=1 N=8

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

NOTE: F Statistic for Roy's Greatest Root is an upper bound.
NOTE: F Statistic for Wilks' Lambda is exact.

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Multivariate Test: Onlyx1

Error Matrix (E)

0.3521207721	0.1323723295	-0.081582595
0.1323723295	0.3986006993	0.2479046457
-0.081582595	0.2479046457	1.2273557756
-0.238505417	0.2352986084	1.2221331781
-0.211550264	0.2252678955	1.0891872222

Error Matrix (E)

-0.238505417	-0.211550264
0.2352986084	0.2252678955
1.2221331781	1.0891872222
1.4505775203	1.3142693993
1.3142693993	1.2243378999

Hypothesis Matrix (H)

0.9444331687	2.1396606906	2.4004527379
2.1396606906	4.8475085618	5.4383460192
2.4004527379	5.4383460192	6.1011975631
2.3189700315	5.2537428628	5.8940940938
2.2097384417	5.0062732199	5.6164616708

Hypothesis Matrix (H)

2.3189700315	2.2097384417
5.2537428628	5.0062732199
5.8940940938	5.6164616708
5.6940206946	5.4258124278
5.4258124278	5.1702377073

'Output 3.1'
'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Multivariate Test: Onlyx1

Multivariate Statistics and Exact F Statistics

S=1 M=1.5 N=8

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.06558545	51.29	5	18
Pillai's Trace	0.93441455	51.29	5	18
Hotelling-Lawley Trace	14.24728504	51.29	5	18
Roy's Greatest Root	14.24728504	51.29	5	18

Multivariate Statistics and Exact F Statistics

S=1 M=1.5 N=8

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

'Multivariate Regression for Fish Data'

The REG Procedure
Model: MODEL1
Multivariate Test: Onlyx2

Error Matrix (E)

0.3521207721	0.1323723295	-0.081582595
0.1323723295	0.3986006993	0.2479046457
-0.081582595	0.2479046457	1.2273557756
-0.238505417	0.2352986084	1.2221331781
-0.211550264	0.2252678955	1.0891872222

Error Matrix (E)

-0.238505417	-0.211550264
0.2352986084	0.2252678955
1.2221331781	1.0891872222
1.4505775203	1.3142693993
1.3142693993	1.2243378999

Hypothesis Matrix (H)

0.0976730369	0.0792504059	-0.01314736
0.0792504059	0.0643025654	-0.010667567
-0.01314736	-0.010667567	0.0017697113
-0.020670718	-0.016771904	0.0027823991
-0.032603961	-0.026454355	0.0043886832

Hypothesis Matrix (H)

-0.020670718	-0.032603961
-0.016771904	-0.026454355
0.0027823991	0.0043886832
0.0043745806	0.0069000341
0.0069000341	0.0108834365

```
                'Multivariate Regression for Fish Data'

                         The REG Procedure
                           Model: MODEL1
                    Multivariate Test: Onlyx2

          Multivariate Statistics and Exact F Statistics

                       S=1     M=1.5     N=8

Statistic                         Value   F Value   Num DF   Den DF

Wilks' Lambda                0.65490563      1.90        5       18
Pillai's Trace               0.34509437      1.90        5       18
Hotelling-Lawley Trace       0.52693756      1.90        5       18
Roy's Greatest Root          0.52693756      1.90        5       18

          Multivariate Statistics and Exact F Statistics

                       S=1     M=1.5     N=8

                Statistic                      Pr > F

                Wilks' Lambda                  0.1449
                Pillai's Trace                 0.1449
                Hotelling-Lawley Trace         0.1449
                Roy's Greatest Root            0.1449
```

توجه کنید که تمامی خروجی های ادامه ۳-۱ مربوط بـه PROC REG مـی باشـد.یکـی از خاصیت های جالب این روش این است که می توان برای آزمون فرض هایی که بطـور مثـال بصورت $H_0: \beta_1 - \beta_2 = c_0$ باشد نیز میتـوان مقـادیر را محاسـبه کـرد کـه دسـتور آن بصورت :

```
mtest x1-x2=c0/print;
```

است .برای مثال اگر $C_0 = 2$ قرار دهیم،خروجی برنامه بصورت زیر است :

ادامه خروجی ۳–۱– آزمون $H_0: \ \beta_1 - \beta_2 = c_0$

The REG Procedure
Model: MODEL1
Multivariate Test: inter2

Error Matrix (E)

0.3521207721	0.1323723295	-0.081582595
0.1323723295	0.3986006993	0.2479046457
-0.081582595	0.2479046457	1.2273557756
-0.238505417	0.2352986084	1.2221331781
-0.211550264	0.2252678955	1.0891872222

Error Matrix (E)

-0.238505417	-0.211550264
0.2352986084	0.2252678955
1.2221331781	1.0891872222
1.4505775203	1.3142693993
1.3142693993	1.2243378999

Hypothesis Matrix (H)

0.6803866523	0.4902658391	0.6639113665
0.4902658391	0.3532705883	0.4783942513
0.6639113665	0.4783942513	0.647835023
0.6982429954	0.5031325745	0.6813353255
0.7487641599	0.5395365824	0.7306331405

Hypothesis Matrix (H)

0.6982429954	0.7487641599
0.5031325745	0.5395365824
0.6813353255	0.7306331405
0.7165679676	0.76841503
0.76841503	0.8240134711

```
                    'Output 3.1'
        'Multivariate Regression for Fish Data'

                 The REG Procedure
                   Model: MODEL1
             Multivariate Test: inter2

      Multivariate Statistics and Exact F Statistics

                 S=1    M=1.5    N=8

Statistic                    Value   F Value   Num DF   Den DF

Wilks' Lambda           0.18915209     15.43        5       18
Pillai's Trace          0.81084791     15.43        5       18
Hotelling-Lawley Trace  4.28675105     15.43        5       18
Roy's Greatest Root     4.28675105     15.43        5       18

       Multivariate Statistics and Exact F Statistics

                 S=1    M=1.5    N=8

                Statistic                Pr > F

                Wilks' Lambda            <.0001
                Pillai's Trace           <.0001
                Hotelling-Lawley Trace   <.0001
                Roy's Greatest Root      <.0001
```

۳-۲- محاسبات مربوط به روش Stepdown

مثال ۳-۲-

در این مثال ما با استفاده داده های مثال ماهی می خواهیم یکی از روش های انتخاب مدل های رگرسیونی را بنا به نیاز مسئله،با کمک نرم افزار SAS انجام دهیم.

با توجه به روش Stepdown در این مثال می خواهیم تاثیر نسبت ماهی های مرده را در ساعت های اندازه گیری شده به شرط ساعت های قبلی را بررسی کنیم . و در آخر با پذیریش همه آنها مدل بصورت :

$$H_0 = H_{01} \bigcap H_{02} \bigcap H_{03} \bigcap H_{04} \bigcap H_{05},$$

رابطه (۴۷)

در می آید که در آن داریم :

$\mathbf{H_{01}}: x_1$ و x_2 هیچ گونه تاثیری روی y_1 ندارند

$\mathbf{H_{02}}: x_1$ و x_2 هیچ گونه تاثیری روی y_2 به شرط y_1 ندارند

$\mathbf{H_{03}}: x_1$ و x_2 هیچ گونه تاثیری روی y_3 به شرط y_1 و y_2 ندارند

$\mathbf{H_{04}}: x_1$ و x_2 هیچ گونه تاثیری روی y_4 به شرط y_1 ، y_2 و y_3 ندارند

$\mathbf{H_{05}}: x_1$ و x_2 هیچ گونه تاثیری روی y_5 به شرط y_1 ، y_2 ، y_3 و y_4 ندارند

ما در اینجا بترتیب تمامی ۵ فرضیه بالا را انجام میدهیم و اگر بخواهیم سطح معنی داری کلی برابر $\alpha = 0.05$ باشد یعنی $(1-\alpha_1)(1-\alpha_2)\cdots(1-\alpha_p) = (1-\alpha)$، می توان

α^* را بصورت $\alpha^* = 1-(1-0.05)^{\frac{1}{5}} \cong 0.01$ در نظر گرفت و تمامی آزمـون فـرض هـا را در سطح 0.01 **انجام میدهیم.**

همچنین این توضیح در اینجا لازم است که منظور از آزمون فرض های شرطی این است که مثلا در آزمون فرض H_{03} ما مقادیر متغیر های وابسته y1 **و** y2 **را بعنوان متغیر هـای مستقل (ثابت) در کنار متغیر های** x1 **و** x2 **در نظر می گیریم ، یعنی :**

$$y_3 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \gamma_1 y_1 + \gamma_2 y_2 + \epsilon_3. \quad \text{رابطه (۴۸)}$$

و آزمون ما بصورت :

$$H_{03}:\ \beta_1 = \beta_2 = 0$$

که محاسبات مربوط به آن با استفاده از دستور زیر انجام می شود :

```
proc reg;
model y3 = x1 x2 y1 y2;
test x1 = 0.0, x2=0.0;
```

رابطه (۴۹)

برنامه ۳–۲–

```
/*program 3-2*/
options ls=64 ps=45 nodate nonumber;
title1 'Output 3.4 ';
title2 ' Stepdown Analysis';
/*The following program performs the Stepdown Analysis */
proc reg data=fish;
model y1=x1 x2;
fishwt: test x2=0.0;
fmodel: test x1=0.0,x2=0.0;
proc reg data=fish;
model y2=x1 x2 y1;
fishwt: test x2=0.0;
fmodel: test x1=0.0,x2=0.0;
proc reg data=fish;
model y3=x1 x2 y1 y2;
fishwt: test x2=0.0;
fmodel: test x1=0.0,x2=0.0;
proc reg data=fish;
model y4=x1 x2 y1 y2 y3;
fishwt: test x2=0.0;
fmodel: test x1=0.0,x2=0.0;
proc reg data=fish;
model y5=x1 x2 y1 y2 y3 y4;
fishwt: test x2=0.0;
fmodel: test x1=0.0,x2=0.0;
run;
```

خروجی ۳–۲–

'Output 3.4 '
' Stepdown Analysis'

The REG Procedure
Model: MODEL1
Dependent Variable: y1

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	0.95388	0.47694	29.80
Error	22	0.35212	0.01601	
Corrected Total	24	1.30600		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.12651	R-Square	0.7304
Dependent Mean	0.19092	Adj R-Sq	0.7059
Coeff Var	66.26628		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-1.65110	0.28363	-5.82	<.0001
x1	1	0.33640	0.04379	7.68	<.0001
x2	1	-0.43079	0.17439	-2.47	0.0217

'Output 3.4 '
' Stepdown Analysis'

The REG Procedure
Model: MODEL1
Dependent Variable: y2

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	3	4.95786	1.65262	99.49
Error	21	0.34884	0.01661	
Corrected Total	24	5.30670		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.12888	R-Square	0.9343
Dependent Mean	0.64051	Adj R-Sq	0.9249
Coeff Var	20.12236		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-3.37918	0.46054	-7.34	<.0001
x1	1	0.63568	0.08561	7.43	<.0001
x2	1	-0.18759	0.20079	-0.93	0.3608
y1	1	0.37593	0.21720	1.73	0.0982

' Stepdown Analysis'

The REG Procedure
Model: MODEL1
Dependent Variable: y3

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	4	6.71616	1.67904	34.06
Error	20	0.98599	0.04930	
Corrected Total	24	7.70215		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.22204	R-Square	0.8720
Dependent Mean	1.02471	Adj R-Sq	0.8464
Coeff Var	21.66803		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-2.20097	1.49774	-1.47	0.1572
x1	1	0.42534	0.28082	1.51	0.1455
x2	1	0.10798	0.35302	0.31	0.7629
y1	1	-0.53190	0.39997	-1.33	0.1985
y2	1	0.79858	0.37593	2.12	0.0463

'Stepdown Analysis'

The REG Procedure
Model: MODEL1
Dependent Variable: y4

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	5	7.37355	1.47471	185.20
Error	19	0.15130	0.00796	
Corrected Total	24	7.52485		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.08924	R-Square	0.9799
Dependent Mean	1.08543	Adj R-Sq	0.9746
Coeff Var	8.22124		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.23937	0.63361	-0.38	0.7098
x1	1	0.07028	0.11916	0.59	0.5622
x2	1	-0.12525	0.14221	-0.88	0.3895
y1	1	-0.53805	0.16771	-3.21	0.0046
y2	1	0.19665	0.16726	1.18	0.2542
y3	1	0.92026	0.08987	10.24	<.0001

'Output 3.4 '
' Stepdown Analysis'

The REG Procedure
Model: MODEL1
Dependent Variable: y5

Number of Observations Read	25
Number of Observations Used	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	6	6.76074	1.12679	660.66
Error	18	0.03070	0.00171	
Corrected Total	24	6.79144		

Analysis of Variance

Source	Pr > F
Model	<.0001
Error	
Corrected Total	

Root MSE	0.04130	R-Square	0.9955
Dependent Mean	1.10294	Adj R-Sq	0.9940
Coeff Var	3.74441		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.08338	0.29433	-0.28	0.7802
x1	1	0.01863	0.05565	0.33	0.7417
x2	1	0.08896	0.06715	1.32	0.2018
y1	1	0.04755	0.09637	0.49	0.6277
y2	1	0.02851	0.08018	0.36	0.7263
y3	1	-0.12745	0.10619	-1.20	0.2456

'Output 3.4 '
' Stepdown Analysis'

The REG Procedure
Model: MODEL1
Dependent Variable: y5

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
y4	1	1.01661	0.10617	9.57	<.0001

پس در تمامی خروجی های بالا ما مقدار $\alpha^* \cong 0.01$ گرفته و آزمون های بالا را با توجه به مقدار α و P-value تحلیل می کنیم :

آزمون ۱ : H_{01} (x1 وx2 هیچ گونه تاثیری روی y_1 ندارند) که با توجه به خروجی ۳-۲ دیده می شود که فرض H_{01} رد می شود یعنی پارامتر ثابت ومتغیر x_1 در مدل تاثیر دارد(با توجه به خروجی ۳-۲)

آزمون ۲ : H_{02} (x_1 وx_2 هیچ گونه تاثیری روی y_2 به شرط y_1 ندارند)

فرضیه H_{02} رد می شود و با توجه به خروجی باز پارامتر ثابت و متغیر x_1 در مدل ما تاثیر دارند.

آزمون ۳ : H_{03}(x_1 وx_2 هیچ گونه تاثیری روی y_3 به شرط y_1 و y_2 ندارند)

۲ فرضیه H_{03} رد می شود ، یعنی مدل ما متاثر از متغیر y_2 است.

آزمون ۴ : H_{04} (x_1 وx_2 هیچ گونه تاثیری روی y_4 به شرط y_1 ، y_2 و y_3 ندارند)

فرضیه H_{04} رد می شود ، یعنی مدل ما متاثر از متغیر y_3 است

آزمون ۵ : H_{05} (x_1 وx_2 هیچ گونه تاثیری روی y_5 به شرط y_1 ، y_2 ، y_3 و y_4 ندارند)

فرض H_{05} رد می شود،یعنی با توجه به خروجی متغیر y_4 در مدل ما تاثیر دارد.

۳-۳-آزمون نرمال بودن خطاها

یکی از فرض های اساسی که در رگرسیون چند متغیره در نظر گرفته می شود ایــن است که بردار خطاها از توزیع نرمال p متغیره آمده اند. برای انجام آزمون اینکه خطا ها از این توزیع آمده اند یا خیر،می توان از نمودار q-q plot استفاده کرد.

برای این منظور ما در مثال ۳-۳ با ارائه کردن برنامه مربوط به آن در نرم افزار SAS ،نمودار q-q plot را رسم می کنیم.

مثال ۳-۳-

داده های این مثال از کتاب Dr.William D. Rohwer reported in Timm (1975, p. 281, 345) آورده شده است. در این مثال می خواهیم ۳ آزمون مختلـف(متغیــر وابسته) را برای ۳۲ مدرسه برحسب متغیر های توضیحی(معیارهایی می باشــند بــرای ۵ نــوع جفت مرتبط با یکدیگر (نمره از ۲۰)) که در زیر تعریف شده اند پیش بینی کنیم .

Peabody Picture Vocabulary Test (y_1)
Raven Progressive Matrices Test (y_2)
Student Achievement Test (y_3)
named (x_1)
still (x_2)
named still (x_3)
named action (x_4)
sentence still(x_5)

داده های این مسئله بصورت زیر می باشند :

OBS	PPVT	RPMT	SAT	N	S	NS	NA	SS
1	68	15	24	0	10	8	21	22
2	82	11	8	7	3	21	28	21
3	82	13	88	7	9	17	31	30
4	91	18	82	6	11	16	27	25
5	82	13	90	20	7	21	28	16
6	100	15	77	4	11	18	32	29
7	100	13	58	6	7	17	26	23
8	96	12	14	5	2	11	22	23
9	63	10	1	3	5	14	24	20
10	91	18	98	16	12	16	27	30
11	87	10	8	5	3	17	25	24
12	105	21	88	2	11	10	26	22
13	87	14	4	1	4	14	25	19
14	76	16	14	11	5	18	27	22
15	66	14	38	0	0	3	16	11
16	74	15	4	5	8	11	12	15
17	68	13	64	1	6	10	28	23
18	98	16	88	1	9	12	30	18
19	63	15	14	0	13	13	19	16
20	94	16	99	4	6	14	27	19
21	82	18	50	4	5	16	21	24
22	89	15	36	1	6	15	23	28
23	80	19	88	5	8	14	25	24
24	61	11	14	4	5	11	16	22
25	102	20	24	5	7	17	26	15
26	71	12	24	0	4	8	16	14
27	102	16	24	4	17	21	27	31
28	96	13	50	5	8	20	28	26
29	55	16	8	4	7	19	20	13
30	96	18	98	4	7	10	23	19
31	74	15	98	2	6	14	25	17
32	78	19	50	5	10	18	27	26

جدول ۳

در اینجا می خواهیم مدل چند متغیره رگرسیونی :

$$Y_{32\times3} = X_{32\times6}B_{6\times3} + \mathcal{E}_{32\times3},$$ رابطه (۵۰)

را برازش دهیم،که محاسبات لازم و خروجی های مربوط به تست نرمال بودن در برنامه ۳-۳ و

خروجی ۳-۳ آمده است .

برنامه ۳-۳-

```
options ls=64 ps=45 nodate nonumber;
title1 'Output 3-3';
title2 'Multivariate Regression Diagnostics';
/* Store all the residuals, predicted values and the diagonal
elements (p_ii) of the projection matrix, (X(X'X)^-1X'), in a
SAS data set (we call it B in the following) */
proc reg data=rohwer;
model y1-y3=x1-x5/noprint;
output out=b r=e1 e2 e3 p=yh1 yh2 yh3 h=p_ii;
/* Test for multivariate normality using the skewness and kurtosis
measures of the residuals. Since the sample means of the residuals
are zeros, Program 1.2 essentially works. */
proc iml;
use b;
read all var {e1 e2 e3} into y;
n = nrow(y) ;
p = ncol(y) ;
dfchi = p*(p+1)*(p+2)/6;
q = i(n) - (1/n)*j(n,n,1);
s = (1/(n))*y`*q*y; /* Use the ML estimate of Sigma */
s_inv = inv(s);
g_matrix = q*y*s_inv*y`*q;
beta1hat = ( sum(g_matrix#g_matrix#g_matrix) )/(n*n);
beta2hat =trace( g_matrix#g_matrix )/n;
kappa1 = n*beta1hat/6;
kappa2 = (beta2hat - p*(p+2) ) /sqrt(8*p*(p+2)/n);
pvalskew = 1 - probchi(kappa1,dfchi);
pvalkurt = 2*( 1 - probnorm(abs(kappa2)) ) ;
print beta1hat ;
print kappa1 ;
print pvalskew;
print beta2hat ;
print kappa2 ;
print pvalkurt;
/* Q-Q plot for checking the multivariate normality using
Mahalanobis distance of the residuals;*/
```

```
data b;
set b;
totn=32.0; /* totn is the number of observations */
k=5.0; /* k is the no. of indep. variables */
p=3.0; /* p is the number of dep. variables */
proc princomp data=b cov std out=c noprint;
var e1-e3;
data sqd;
set c;
student=_n_;
dsq=uss(of prin1-prin3);
dsq=dsq*(totn-k-1)/(totn-1);
* Divide the distances by (1-p_ii);
dsq=dsq/(1-p_ii);
```

```
data qqp;
set sqd;
proc sort;
by dsq;
data qqp;
set qqp;
stdnt_rs=student;
chisq=cinv(((_n_-.5)/ totn),p);
proc gplot data=qqp;
plot dsq*chisq;
label dsq = 'Mahalanobis D Square'
chisq= 'Chi-Square Quantile';
run;
* Q-Q plot for detection of outliers using Robust distance;
* (Section 3.10.3);
data qqprd;
set sqd;
rdsq=((totn-k-2)*dsq/(totn-k-1))/(1-(dsq/(totn-k-1)));
proc sort;
by rdsq;
data qqprd;
set qqprd;
stdnt_rd=student;
chisq=cinv(((_n_-.5)/ totn),p);
proc gplot;
plot rdsq*chisq;
label rdsq = 'Robust Mahalanobis D Square'
chisq= 'Chi-Square Quantile';
run;
* Influence Measures for Multivariate Regression;
* (Section 3.10.4);
* Diagonal Elements of Hat (or Projection) Matrix;
data hat;
set sqd;
```

```
proc sort;
by p_ii;
data hat;
set hat;
stdnt_h=student;
keep stdnt_h p_ii;
run;
* Cook's type of distance for detection of influential obs.;
data cookd;
set sqd;
csq=dsq*p_ii/((1-p_ii)*(k+1));
proc sort;
by csq;
data cookd;
set cookd;
stdnt_co=student;
keep stdnt_co csq;
run;
```

```
* Welsch-Kuh type statistic for detection of influential obs.;
data wks;
set qqprd;
wksq=rdsq*p_ii/(1-p_ii);
proc sort;
by wksq;
data wks;
set wks;
stdnt_wk=student;
keep stdnt_wk wksq;
run;
* Covariance Ratio for detection of influential obs.;
data cvr;
set qqprd;
covr=((totn-k-2)/(totn-k-2+rdsq))**(k+1)/(1-p_ii)**p;
covr=covr*((totn-k-1)/(totn-k-2))**((k+1)*p);
proc sort;
by covr;
data cvr;
set cvr;
stdnt_cv=student;
keep stdnt_cv covr;
run;
* Display of Influence Measures;
data display;
merge hat cookd wks cvr;
title2 'Multivariate Regression Diagnostics';
title3 'Influence Measures';
proc print data=display noobs;
var stdnt_h p_ii stdnt_co csq stdnt_wk wksq stdnt_cv covr;
```

خروجی ۳–۳–

'Output 3-3'
'Multivariate Regression Diagnostics'

BETA1HAT
1.1489572

KAPPA1
6.1277716

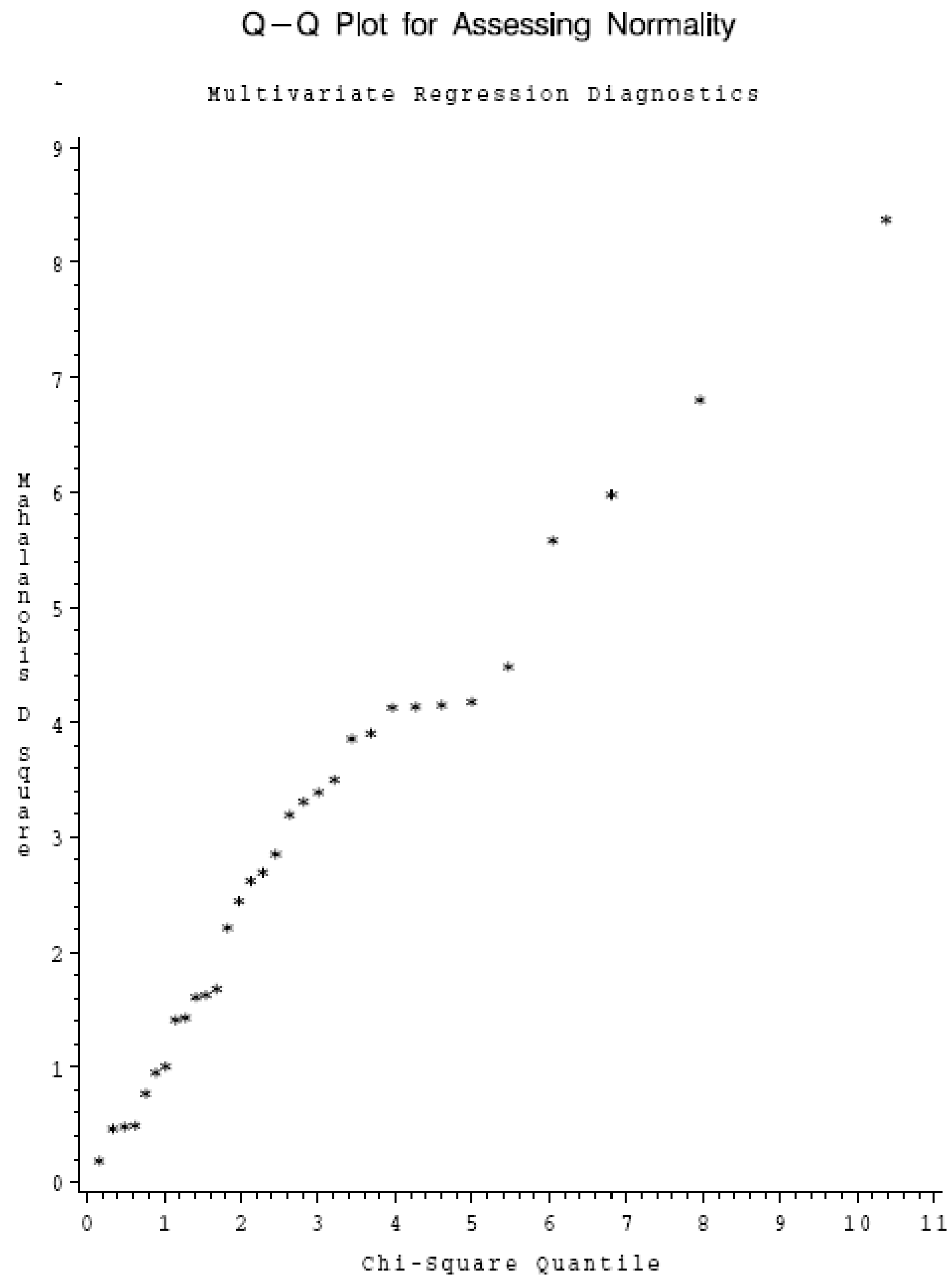
PVALSKEW
0.8044166

BETA2HAT
13.435642

KAPPA2
-0.807831

PVALKURT
0.419188

dnt_h	p_ii	nt_co	csq	nt_wk	wksq	nt_cv	covr
23	0.04455	4	0.00645	4	0.03795	25	0.3287
7	0.04531	6	0.01458	6	0.08572	14	0.4920
9	0.05131	15	0.01519	15	0.08826	31	0.5594
4	0.07314	7	0.02530	7	0.16640	21	0.7485
32	0.07321	23	0.03036	23	0.20610	9	0.7624
20	0.07881	28	0.03422	28	0.21065	23	0.8746
30	0.08655	2	0.03576	2	0.21258	32	0.8891
31	0.08922	20	0.03733	20	0.23950	7	1.0593
13	0.10375	22	0.04025	22	0.24151	17	1.2668
28	0.11190	9	0.04040	26	0.25993	8	1.2851
14	0.12650	26	0.04261	13	0.26907	30	1.3235
21	0.14024	13	0.04267	11	0.28094	20	1.3707
3	0.14173	30	0.04505	9	0.28165	12	1.5332
11	0.14543	11	0.04568	30	0.29195	1	1.5475
26	0.15334	32	0.05503	18	0.33957	13	1.6506
6	0.15432	18	0.05671	10	0.37234	3	1.6629
25	0.15713	10	0.06339	32	0.37831	28	1.9612
1	0.16701	24	0.07294	24	0.44995	29	2.1183
12	0.17050	3	0.07411	3	0.47696	11	2.2117
17	0.17321	31	0.09758	16	0.72235	4	2.2694
8	0.17661	1	0.11067	31	0.73104	26	2.3879
22	0.19380	12	0.11629	1	0.73167	24	2.7132
24	0.20642	16	0.11832	12	0.77168	6	2.9955
2	0.21845	17	0.14448	17	0.99133	22	3.0522
18	0.26354	8	0.14768	8	1.01291	19	3.2436
19	0.29836	21	0.15164	19	1.10294	27	3.3582
29	0.30427	14	0.16427	21	1.11383	2	3.5453
16	0.33183	19	0.17321	14	1.28379	18	4.0558
15	0.33247	25	0.26008	29	2.07746	16	4.8372
27	0.36726	29	0.30260	25	2.21298	15	6.5280
10	0.45161	27	0.33866	27	2.25782	5	9.5932
5	0.56821	5	0.84672	5	5.73670	10	11.0314



همانطور که در خروجی این برنامه می بینید ، مقدار p-value برای آماره چولگی (PVALSKEW) برابر ۰٫۸۰۴۴ و برای معیار کشیدگی توزیع مقدار p-value برابر ۰٫۴۱۹۲ شده است(PVALKURT).لذا هر دو فرض عدم چولگی و عدم کشیدگی پذیرفته می شود.

با استفاده از نمودار Q_Q plot می توان گفت که داده های مربوطه دارای خطای تقریبا نرمال هستند. پس بطور کلی هیچ گونه مدرکی دال بر رد فرض نرمال بودن خطاها موجود نیست.

۳-۴- آزمون همگونی واریانس ها

همانطور که می دانیم در رگرسیون خطی یک متغیره با رسم نمودار های خطاهای استاندارد در مقابل مقادیر پیش بینی شده و یا مقادیر متغیر های مستقل می توان به درست بودن یا نبودن فرضیاتی از مدل همچون همگون بودن واریانس،پی برد.از این رو در رگرسیون چند متغیره نیز مقادیر زیر را محاسبه و نمودار آنها را با استفاده از داده های مثال ۳-۴ رسم می کنیم. در صورت تصادفی بودن نقاط پی به همگونی واریانس در بردار خطاها می بریم.

$$d_i^{\epsilon} = \sqrt{\frac{\hat{\epsilon}_i' \hat{\Sigma}^{-1} \hat{\epsilon}_i}{1 - p_{ii}}} \qquad \text{and} \qquad d_i^{y} = \sqrt{\frac{\hat{\mathbf{y}}_i' \hat{\Sigma}^{-1} \hat{\mathbf{y}}_i}{p_{ii}}},$$

رابطه (۵۱)

که در آن $\mathbf{P_{ii}}$ مقادیر قطر اصلی ماتریس $\mathbf{x(x'x)^{-1}x'}$ می باشند.

مثال ۳-۴-

در این مثال با استفاده از داده های مربوط به آلودگی هوا در منبع ۴ و با استفاده از برنامه ۳-۴ محاسبات لازم را انجام و نمودار D-D plot را رسم می کنیم.

برنامه ۳–۴–

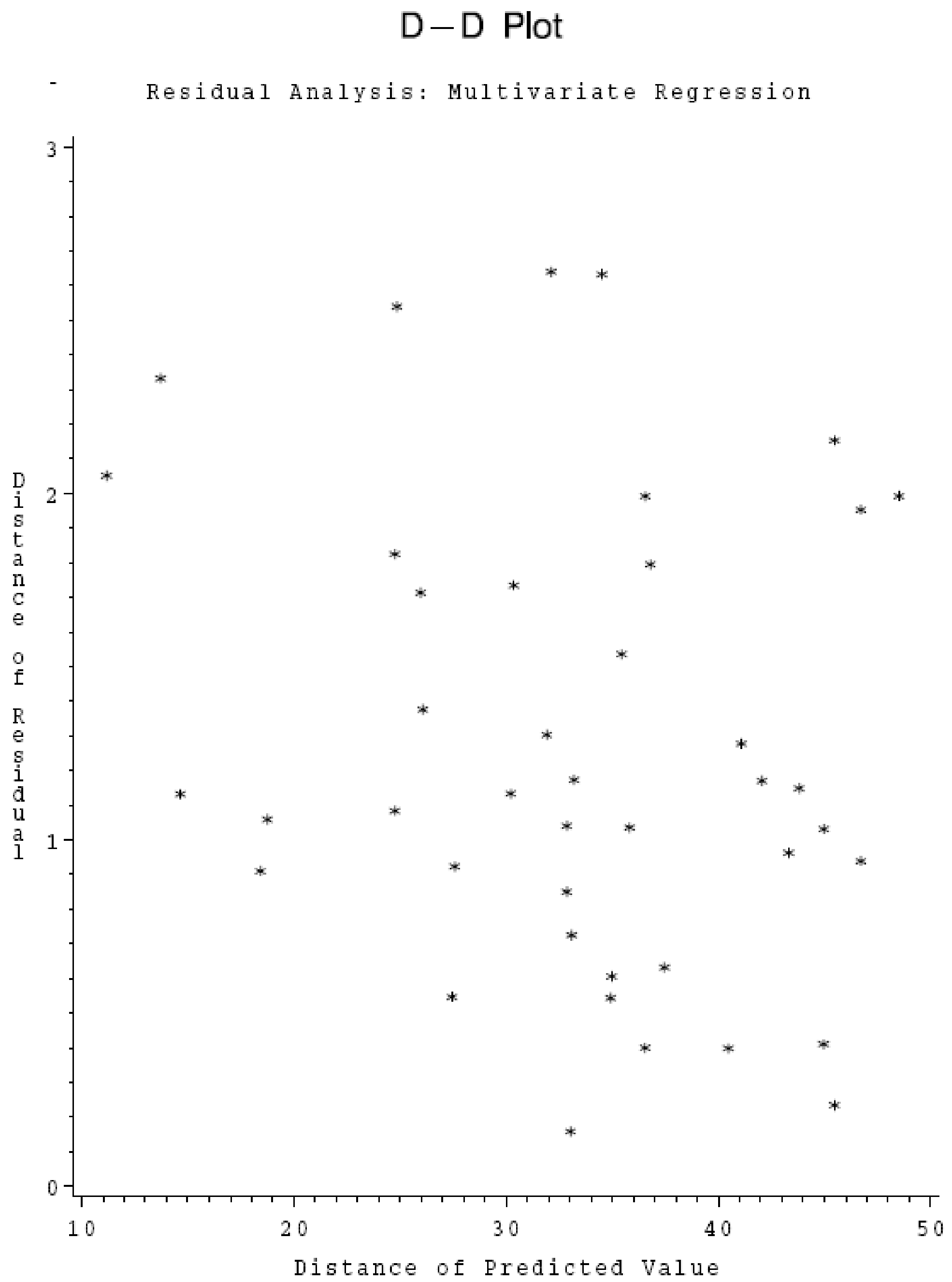
داده های مربوط به این مثال را بصورت زیر وارد برنامه کرده :

```
data pollut;
input x1 x2 y1-y5;
cards;
8 98 7 2 12 8 2
7 107 4 3 9 5 3
7 103 4 3 5 6 3
10 88 5 2 8 15 4
6 91 4 2 8 10 3
8 90 5 2 12 12 4
9 84 7 4 12 15 5
5 72 6 4 21 14 4
7 82 5 1 11 11 3
8 64 5 2 13 9 4
6 71 5 4 10 3 3
6 91 4 2 12 7 3
7 72 7 4 18 10 3
10 70 4 2 11 7 3
10 72 4 1 8 10 3
9 77 4 1 9 10 3
8 76 4 1 7 7 3
8 71 5 3 16 4 4
9 67 4 2 13 2 3
9 69 3 3 9 5 3
10 62 5 3 14 4 4
9 88 4 2 7 6 3
8 80 4 2 13 11 4
5 30 3 3 5 2 3
6 83 5 1 10 23 4
8 84 3 2 7 6 3
6 78 4 2 11 11 3
8 79 2 1 7 10 3
6 62 4 3 9 8 3
10 37 3 1 7 2 3
8 71 4 1 10 7 3
7 52 4 1 12 8 4
5 48 6 5 8 4 3
6 75 4 1 10 24 3
```

```
6 75 4 1 10 24 3
10 35 4 1 6 9 2
8 85 4 1 9 10 2
5 86 3 1 6 12 2
5 86 7 2 13 18 2
7 79 7 4 9 25 3
7 79 5 2 8 6 2
6 68 6 2 11 14 3
8 40 4 3 6 5 2
;
```

```
options ls=64 ps=45 nodate nonumber;
title1 'Output 3.10';
title2 'Multivariate Regression: D-D Plot';
* The data set containing independent and dependent variables.;
* Make transformations if necessary;
data pollut;
set pollut;
x1=log(x1);
108 Applied Multivariate Statistics
x2=log(x2);
y1=log(y1);
y2=log(y2);
y3=log(y3);
y4=log(y4);
y5=log(y5);
proc reg data=pollut;
model y3 y4=x1 x2/noprint;
output out=b r=e1 e2 p=yh1 yh2 h=p_ii;
/* Univariate residual plots;
proc reg data=pollut;
model y3 y4=x1 x2 /noprint;
plot student.*p.;
title2 'Univariate Residual Plots';
run;
*/
```

خروجی ۳-۴-



با استفاده از نمودار بالا تا حدودی متوجه یک روند صعودی می شویم که ممکن است دلیل عدم همگونی واریانس در بردار خطا ها باشد.

۳–۵– فاصله اطمینان همزمان

در آزمون فرض هاممکن است در صورت رد فـرض LB=0 علاقمنـد بـه یـافتن فاصله اطمینان همزمان تکه به تک برای تمامی LB ها باشیم.برای این منظور با فـرض کامـل بودن رتبه ماتریس X ، می توان برای ترکیب خطـی از نـوع $\mathbf{c'LBd}$ بـا اسـتفاده از آزمـون فرض خطی $\mathbf{LB = 0}$ ، یک فاصله اطمینان همزمان ساخت.

توجه داشته باشید که آزمون با نماد $H_0^{(\mathbf{c},\mathbf{d})}\ :\ \mathbf{c'LBd = 0}$ بدین معنـی اسـت کـه فرض بالا برای تمامی بردار های c,d برقرار است.لذا برای ساختن فاصله اطمینـان همزمـان و همچنین آزمون بالا از آماره F استفاده می کنیم؛که در اینجا آن را با نماد $F^{(\mathbf{c},\mathbf{d})}$ نمـایش مـی دهیم .و با توجه به مطالعـات قبلـی در دروس گذشـته فـرض $\mathbf{H_0}$ پذیرفتـه مـی شـود اگـر $\max_{\mathbf{c},\mathbf{d}} F^{(\mathbf{c},\mathbf{d})} \le F_\alpha(h, n-k-h+r-1)$ ، بنابراین فاصله اطمینان همزمان بصورت :

$$\mathbf{c'L\hat{B}d} \pm \{ \frac{h}{n-k-h+r-1} F_\alpha(h, n-k-h+r-1)\mathbf{c'L(X'X)^{-1}L'c}\ \cdot\ \mathbf{d'Ed}\}^{1/2}$$

حال برای بررسی بیشتر با آوردن یک مثال طرز بدست آوردن فاصـله اطمینـان همزمـان را در SAS با استفاده از برنامه ۳–۵ توضیح می دهیم.

مثال ۳–۵–

ما در این مثال می خواهیم برای داده های مربوط به چوب پنبـه درختـان[1] فاصـله اطمینان همزمان برای $\mu_1-\mu_2+\mu_3-\mu_4$ ، $\mu_3-\mu_4$ و $\mu_1-\mu_3$ و ما در اینجـا بـا قـرار دان بردار d بصورت

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 0 \\ 1 & 1 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

و همچنـین $\mathbf{c} = 1, \mathbf{L} = 1, \text{ and } \mathbf{B} = (\mu_1\ \mu_2\ \mu_3\ \mu_4)$ برنامـه مربوطـه را بـصورت زیـر داریم:

[1] Rao(1948)

برنامه ۳-۵-

```
options ls=64 ps=45 nodate nonumber;
title1 'Output 3.5';
proc iml;
alpha = .05 ;
n = 28;
/*
Calculations for simultaneous confidence intervals are shown below.
r_t =Rank (H+E)=p=4, Rank(L)=1, k=0
df of error matrix = dferror = 27
s = min(r, r_t) = 1, h=max(r,r_t)=4
m1 = .5(|r-r_t| - 1) = 1
m2 = .5(n-k-r_t-2) = 11.
lambda = [h/(n-k-h+r-1)]F(alpha,h, n-k-h+r-1)
*/
r_t=4;
r=1;
k=0;
s = min(r, r_t);
h=max(r,r_t);
m1 = .5*(abs(r-r_t) - 1);
m2 = .5*(n-k-r_t-2);
lambda = (h/(n-k-h+r-1))*finv(1-alpha,h,n-k-h+r-1);
cutoff = sqrt(lambda);
/*
Cut-off point for Bonferroni intervals will be computed as follows:
dferror = 27 ; * dferror=n-1;
g=3.0; * g is the no. of comparisons;
cutoff = tinv(1-(alpha/(2*g)),dferror);
cutoff=cutoff/sqrt(dferror);
*/

xpx = {28};
e = {7840.9643 6041.3214 7787.8214 6109.3214,
6041.3214 5938.1071 6184.6071 4627.1071,
7787.8214 6184.6071 9450.1071 7007.6071,
6109.3214 4627.1071 7007.6071 6102.1071};
bhat = {50.535714 46.178571 49.678571 45.178571};
l = {1} ;
c={1};
d1={1,-1,1,-1};
d2={0,0,1,-1};
d3={1,0,-1,0};
clbhatd1=c`*l*bhat*d1;
clbhatd2=c`*l*bhat*d2;
clbhatd3=c`*l*bhat*d3;
```

```
cwidth1=cutoff*sqrt((c`*l*(inv(xpx))*l*c)*(d1`*e*d1));
cwidth2=cutoff*sqrt((c`*l*(inv(xpx))*l*c)*(d2`*e*d2));
cwidth3=cutoff*sqrt((c`*l*(inv(xpx))*l*c)*(d3`*e*d3));
cl11=clbhatd1-cwidth1;
cl12=clbhatd1+cwidth1;
cl21=clbhatd2-cwidth2;
cl22=clbhatd2+cwidth2;
cl31=clbhatd3-cwidth3;
cl32=clbhatd3+cwidth3;
print 'Simultaneous Confidence Intervals';
print 'For first contrast: (' cl11', ' cl12 ')';
print 'For second contrast:(' cl21', ' cl22' )';
print 'For third contrast: (' cl31', ' cl32 ')';
run;
```

که در آن مقادیر متغیر ها در برنامه توضیح داده شده است . ومقادیر ماتریس $\hat{\mathbf{B}},(\mathbf{X}'\mathbf{X})^{-1},\mathbf{E}$ از برنامه PROC REG و PROC GLM محاسبه شده است(با توجه به اینکه خروجی های مانند آن برای مثال ۳-۱ آمده است، از آوردن خروجی های مربوط به این قسمت بدلیل تکـرار صرف نظر می کنیم).

خروجی ۳-۵-

خروجی مربوط به برنامه بالا بصورت زیر می باشد :

```
'Output 3.5'
Simultaneous Confidence Intervals

                                CL11          CL12
For first contrast: ( 1.2786666 ,  16.435619 )

                               CL21          CL22
For second contrast:( -0.539816 ,  9.5398157  )

                                CL31          CL32
For third contrast: ( -4.467176 ,  6.1814617 )
```

در پایان امید است که این پروژه گامی هر چند کوچک در جهت یادگیری هر چه بهتر و عمیق تر مفاهیم و روش های آماری و استفاده بهینه از آنها باشد .

پیوست

پیوست

A) آماره لاندای ویلکس[1]

با توجه به اینکه ما در روش های چند متغیره پیوسته تمامی آزمون های مربوط به میانگین های چند متغیری را با استفاده از آماره F خواندیم ؛ لذا ما در قسمت توضیحی مختصر از آماره ویلکس و تبدیل این آماره به آماره F را در جدولی آورده ایم.

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|}$$

که این آماره برای آزمون فرض $H_0\colon \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$ استفاده می شود .

و اگر $\Lambda \leq \Lambda_{\alpha,p,q,n-q-1}$ آزمون مربوطه رد می شود .

که مقادیر بحرانی آن در جدول a آمده است.

جدول a)مقادیر پایین ترین سطح بحرانی برای آماره لانـدای ویلکس در سطح $\alpha = 0.05$

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|} = \prod_{i=1}^{s} \frac{1}{1+\lambda_i},$$

[1] برای توضیحات بیشتر به مرجع ۲ بخش ۶,۱,۳ مراجعه کنید

که در آن $\lambda_1, \lambda_2, \ldots, \lambda_s$ مقادیر ویژه ماتریس $\mathbf{E}^{-1}\mathbf{H}$ می باشد .و همچنین توان ^{a} یعنی مقدار را در $۱۰^{-۳}$ ضرب کنید. و فرض H_0 رد می شود اگر لاندا کمتر از مقادیر جدول باشد .

	ν_H											
ν_E	1	2	3	4	5	6	7	8	9	10	11	12
						$p = 1$						
1	6.16^a	2.50^a	1.54^a	1.11^a	$.868^a$	$.712^a$	$.603^a$	$.523^a$	$.462^a$	$.413^a$	$.374^a$	$.341^a$
2	.098	.050	.034	.025	.020	.017	.015	.013	.011	.010	9.28^a	8.51^a
3	.229	.136	.097	.076	.062	.053	.046	.041	.036	.033	.030	.028
4	.342	.224	.168	.135	.113	.098	.086	.076	.069	.063	.058	.053
5	.431	.302	.236	.194	.165	.144	.128	.115	.104	.096	.088	.082
6	.501	.368	.296	.249	.215	.189	.169	.153	.140	.129	.119	.111
7	.556	.425	.349	.298	.261	.232	.209	.190	.175	.161	.150	.140
8	.601	.473	.396	.343	.303	.271	.246	.225	.208	.193	.180	.169
9	.638	.514	.437	.382	.341	.308	.281	.258	.239	.223	.209	.196
10	.668	.549	.473	.418	.376	.341	.313	.289	.269	.251	.236	.222
11	.694	.580	.505	.450	.407	.372	.343	.318	.297	.278	.262	.247
12	.717	.607	.534	.479	.436	.400	.370	.345	.323	.304	.286	.271
13	.736	.631	.560	.506	.462	.426	.396	.370	.347	.327	.310	.294
14	.753	.652	.583	.529	.486	.450	.420	.393	.370	.350	.332	.315
15	.768	.671	.603	.551	.508	.473	.442	.415	.392	.371	.352	.336
16	.781	.688	.622	.571	.529	.493	.462	.436	.412	.391	.372	.355
17	.792	.703	.639	.589	.548	.512	.482	.455	.431	.410	.390	.373
18	.803	.717	.655	.606	.565	.530	.499	.473	.449	.427	.408	.390
19	.813	.730	.669	.621	.581	.546	.516	.490	.466	.444	.425	.407
20	.821	.741	.683	.636	.596	.562	.532	.505	.482	.460	.440	.423
21	.829	.752	.695	.649	.610	.576	.547	.520	.497	.475	.455	.437
22	.836	.762	.706	.661	.623	.590	.561	.534	.511	.489	.470	.452
23	.843	.771	.717	.673	.635	.603	.574	.548	.524	.503	.483	.465
24	.849	.779	.727	.684	.647	.615	.586	.560	.537	.516	.496	.478
25	.855	.787	.736	.694	.658	.626	.598	.572	.549	.528	.508	.490
26	.860	.794	.744	.703	.668	.637	.609	.583	.560	.539	.520	.502
27	.865	.801	.752	.712	.677	.647	.619	.594	.571	.551	.531	.513
28	.870	.807	.760	.721	.686	.656	.629	.604	.582	.561	.542	.524
29	.874	.813	.767	.729	.695	.665	.638	.614	.592	.571	.552	.535
30	.878	.819	.774	.736	.703	.674	.647	.623	.601	.581	.562	.544
40	.907	.861	.824	.793	.766	.741	.718	.696	.677	.658	.641	.625
60	.938	.905	.879	.856	.835	.816	.798	.781	.766	.751	.736	.723
80	.953	.928	.907	.889	.873	.858	.843	.829	.816	.804	.792	.780
100	.962	.942	.925	.910	.897	.884	.872	.860	.849	.838	.828	.818
120	.968	.951	.937	.925	.913	.902	.891	.882	.872	.863	.854	.845
140	.973	.958	.946	.935	.925	.915	.906	.897	.889	.881	.873	.865
170	.978	.965	.955	.946	.937	.929	.922	.914	.907	.900	.893	.887
200	.981	.970	.962	.954	.947	.940	.933	.926	.920	.914	.908	.902
240	.984	.975	.968	.961	.955	.949	.944	.938	.933	.928	.923	.918
320	.988	.981	.976	.971	.966	.962	.957	.953	.949	.945	.941	.937
440	.991	.986	.982	.979	.975	.972	.969	.966	.963	.960	.957	.954
600	.994	.990	.987	.984	.982	.979	.977	.975	.972	.970	.968	.966
800	.995	.993	.990	.988	.986	.984	.983	.981	.979	.977	.976	.974
1000	.996	.994	.992	.991	.989	.988	.986	.985	.983	.982	.981	.979

(continued)

ν_E	ν_H 1	2	3	4	5	6	7	8	9	10	11	12
						$p = 2$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	2.50[a]	.641[a]	.287[a]	.162[a]	.104[a]	.072[a]	.053[a]	.041[a]	.032[a]	.026[a]	.022[a]	.018[a]
3	.050	.018	9.53[a]	5.84[a]	3.95[a]	2.85[a]	2.15[a]	1.68[a]	1.35[a]	1.11[a]	.928[a]	.787[a]
4	.136	.062	.036	.023	.017	.012	9.56[a]	7.62[a]	6.21[a]	5.17[a]	4.36[a]	3.73[a]
5	.224	.117	.074	.051	.037	.028	.023	.018	.015	.013	.011	.009
6	.302	.175	.116	.084	.063	.049	.040	.033	.027	.023	.020	.017
7	.368	.230	.160	.119	.092	.074	.060	.050	.042	.036	.032	.028
8	.4256	.280	.203	.155	.122	.099	.082	.069	.059	.051	.045	.040
9	.473	.326	.243	.190	.153	.126	.106	.090	.078	.068	.060	.053
10	.514	.367	.281	.223	.183	.152	.129	.111	.097	.085	.075	.067
11	.549	.404	.316	.255	.212	.179	.153	.133	.116	.102	.091	.082
12	.580	.437	.348	.286	.240	.204	.176	.154	.136	.120	.108	.097
13	.607	.467	.378	.314	.266	.229	.199	.175	.155	.138	.124	.112
14	.631	.495	.405	.340	.291	.252	.221	.195	.174	.156	.141	.128
15	.652	.519	.431	.365	.315	.275	.242	.215	.193	.174	.157	.143
16	.671	.542	.454	.389	.337	.296	.263	.235	.211	.191	.174	.159
17	.688	.562	.476	.410	.359	.317	.282	.254	.229	.208	.190	.174
18	.703	.581	.496	.431	.379	.337	.301	.272	.246	.225	.206	.189
19	.717	.598	.515	.450	.398	.355	.320	.289	.263	.241	.221	.204
20	.730	.614	.532	.468	.416	.373	.337	.306	.279	.256	.236	.218
21	.741	.629	.548	.485	.433	.390	.354	.322	.295	.271	.251	.232
22	.752	.643	.564	.501	.449	.406	.370	.338	.310	.286	.265	.246
23	.762	.656	.578	.516	.465	.422	.385	.353	.325	.300	.279	.259
24	.771	.668	.591	.530	.479	.436	.399	.367	.339	.314	.292	.272
25	.779	.679	.604	.544	.493	.450	.413	.381	.353	.328	.305	.285
26	.787	.689	.616	.556	.506	.464	.427	.395	.366	.341	.318	.297
27	.794	.699	.627	.568	.519	.477	.440	.407	.379	.353	.330	.309
28	.801	.708	.638	.580	.531	.489	.452	.420	.391	.365	.342	.321
29	.807	.717	.648	.591	.542	.501	.464	.432	.403	.377	.354	.332
30	.813	.725	.657	.601	.553	.512	.475	.443	.414	.388	.365	.344
40	.858	.786	.730	.682	.640	.602	.568	.537	.509	.484	.460	.439
60	.903	.853	.811	.774	.741	.710	.682	.656	.632	.609	.588	.568
80	.927	.888	.854	.825	.798	.772	.749	.727	.706	.686	.667	.649
100	.941	.909	.882	.857	.834	.813	.793	.774	.755	.738	.721	.705
120	.951	.924	.900	.879	.860	.841	.823	.807	.791	.775	.760	.746
140	.958	.934	.914	.895	.878	.862	.846	.831	.817	.803	.790	.777
170	.965	.946	.929	.913	.898	.885	.871	.859	.846	.834	.823	.812
200	.970	.954	.939	.926	.913	.901	.889	.878	.867	.857	.847	.837
240	.975	.961	.949	.938	.927	.917	.907	.897	.888	.879	.870	.862
320	.981	.971	.962	.953	.945	.937	.929	.922	.914	.907	.901	.894
440	.986	.979	.972	.965	.959	.953	.948	.942	.937	.932	.926	.921
600	.990	.984	.979	.975	.970	.966	.961	.957	.953	.949	.945	.942
800	.993	.988	.984	.981	.977	.974	.971	.968	.965	.962	.959	.956
1000	.994	.991	.987	.985	.982	.979	.977	.974	.972	.969	.967	.964

[a] Multiply entry by 10^{-3}.

(continued)

ν_E	ν_H 1	2	3	4	5	6	7	8	9	10	11	12
						$p = 3$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.000	.000	.000	.000	.000	.001[a]	.002[a]	.004[a]	.005[a]	.008[a]	.010[a]	.013[a]
3	1.70[a]	.354[a]	.179[a]	.127[a]	.105[a]	.095[a]	.091[a]	.090[a]	.091[a]	.092[a]	.095[a]	.098[a]
4	.034	.010	.004	.002	.001	.001	.809[a]	.659[a]	.562[a]	.496[a]	.449[a]	.416[a]
5	.097	.036	.018	.010	6.36[a]	4.37[a]	3.20[a]	2.46[a]	1.97[a]	1.64[a]	1.40[a]	1.22[a]
6	.168	.074	.040	.024	.016	.011	.008	.006	.004	3.94[a]	3.28[a]	2.79[a]
7	.236	.116	.068	.043	.029	.021	.016	.012	9.49[a]	7.67[a]	6.35[a]	5.35[a]
8	.296	.160	.099	.066	.046	.034	.026	.020	.016	.013	.011	9.00[a]
9	.349	.203	.131	.091	.066	.049	.038	.030	.024	.020	.016	.014
10	.396	.243	.164	.117	.086	.066	.052	.041	.034	.028	.023	.020
11	.437	.281	.196	.143	.108	.084	.067	.054	.044	.037	.031	.026
12	.473	.316	.226	.169	.130	.103	.083	.067	.056	.047	.040	.034
13	.505	.348	.255	.194	.152	.122	.099	.082	.068	.058	.049	.042
14	.534	.378	.283	.219	.174	.141	.116	.096	.081	.069	.059	.051
15	.560	.405	.309	.243	.195	.160	.133	.111	.095	.081	.070	.061
16	.583	.431	.334	.266	.216	.179	.149	.127	.108	.093	.081	.071
17	.603	.454	.357	.288	.236	.197	.166	.142	.122	.106	.092	.081
18	.622	.476	.379	.309	.256	.215	.183	.157	.136	.118	.104	.092
19	.639	.496	.399	.329	.275	.233	.199	.172	.149	.131	.115	.102
20	.655	.515	.419	.348	.293	.250	.215	.187	.163	.144	.127	.113
21	.669	.532	.437	.366	.310	.266	.230	.201	.177	.156	.139	.124
22	.683	.548	.454	.383	.327	.282	.246	.215	.190	.169	.150	.135
23	.695	.564	.470	.399	.343	.298	.260	.229	.203	.181	.162	.146
24	.706	.578	.486	.415	.359	.313	.275	.243	.216	.193	.173	.156
25	.717	.591	.500	.430	.374	.327	.289	.256	.229	.205	.185	.167
26	.727	.604	.514	.444	.388	.341	.302	.269	.241	.217	.196	.178
27	.736	.616	.527	.458	.401	.355	.315	.282	.253	.229	.207	.188
28	.744	.627	.540	.471	.415	.368	.328	.294	.265	.240	.218	.199
29	.752	.638	.552	.483	.427	.380	.340	.306	.277	.251	.229	.209
30	.760	.648	.563	.495	.439	.392	.352	.318	.288	.262	.239	.219
40	.816	.724	.651	.591	.539	.494	.454	.419	.387	.359	.334	.311
60	.875	.808	.752	.704	.661	.623	.587	.555	.526	.498	.473	.449
80	.905	.853	.808	.769	.733	.700	.670	.641	.615	.590	.566	.544
100	.924	.881	.844	.810	.780	.751	.725	.700	.676	.654	.632	.612
120	.936	.900	.868	.839	.813	.788	.764	.742	.721	.700	.681	.663
140	.945	.913	.886	.861	.837	.815	.794	.774	.755	.736	.719	.702
170	.955	.928	.905	.884	.864	.845	.827	.809	.792	.776	.761	.746
200	.961	.939	.919	.900	.883	.866	.850	.835	.820	.806	.792	.779
240	.968	.949	.932	.916	.901	.887	.873	.860	.848	.835	.823	.811
320	.976	.961	.948	.936	.925	.914	.903	.893	.883	.873	.864	.854
440	.982	.972	.962	.953	.945	.937	.929	.921	.913	.906	.899	.891
600	.987	.979	.972	.966	.959	.953	.947	.941	.936	.930	.924	.919
800	.990	.984	.979	.974	.969	.965	.960	.956	.951	.947	.943	.939
1000	.992	.987	.983	.979	.975	.972	.968	.964	.961	.957	.954	.950

[a] Multiply entry by 10^{-3}.

(continued)

	ν_H											
ν_E	1	2	3	4	5	6	7	8	9	10	11	12
						$p = 4$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.001[a]	.001[a]	.001[a]	.002[a]	.002[a]	.002[a]	.003[a]
4	1.38[a]	.292[a]	.127[a]	.075[a]	.052[a]	.040[a]	.033[a]	.029[a]	.026[a]	.025[a]	.023[a]	.022[a]
5	.026	6.09[a]	2.31[a]	1.13[a]	.647[a]	.416[a]	.292[a]	.218[a]	.172[a]	.141[a]	.120[a]	.105[a]
6	.076	.024	.010	5.07[a]	2.90[a]	1.82[a]	1.22[a]	.872[a]	.652[a]	.508[a]	.409[a]	.338[a]
7	.135	.051	.024	.013	7.74[a]	4.94[a]	3.34[a]	2.36[a]	1.74[a]	1.33[a]	1.05[a]	.848[a]
8	.194	.084	.043	.025	.015	.010	6.98[a]	4.99[a]	3.70[a]	2.82[a]	2.21[a]	1.77[a]
9	.249	.119	.066	.040	.026	.017	.012	8.91[a]	6.66[a]	5.11[a]	4.01[a]	3.21[a]
10	.298	.155	.091	.057	.038	.027	.019	.014	.011	8.29[a]	6.54[a]	5.25[a]
11	.343	.190	.117	.077	.053	.037	.027	.021	.016	.012	9.84[a]	7.95[a]
12	.382	.223	.143	.097	.068	.049	.037	.028	.022	.017	.014	.011
13	.418	.255	.169	.117	.085	.063	.047	.037	.029	.023	.019	.015
14	.450	.286	.194	.138	.102	.077	.059	.046	.037	.030	.024	.020
15	.479	.314	.219	.159	.119	.091	.071	.056	.045	.037	.030	.025
16	.506	.340	.243	.180	.136	.106	.083	.067	.054	.044	.037	.031
17	.529	.365	.266	.200	.154	.121	.096	.078	.064	.053	.044	.037
18	.551	.389	.288	.219	.171	.136	.109	.089	.074	.061	.051	.044
19	.571	.410	.309	.239	.188	.151	.123	.101	.084	.070	.059	.051
20	.589	.431	.329	.257	.205	.166	.136	.113	.094	.079	.068	.058
21	.606	.450	.348	.275	.221	.181	.149	.124	.105	.089	.076	.065
22	.621	.468	.366	.292	.237	.195	.162	.136	.115	.098	.085	.073
23	.636	.485	.383	.309	.253	.210	.175	.148	.126	.108	.093	.081
24	.649	.501	.399	.325	.268	.224	.188	.160	.137	.118	.102	.089
25	.661	.516	.415	.340	.283	.237	.201	.172	.148	.128	.111	.097
26	.673	.530	.430	.355	.297	.251	.214	.183	.158	.138	.120	.106
27	.684	.544	.444	.369	.311	.264	.226	.195	.169	.147	.129	.114
28	.694	.556	.458	.383	.324	.277	.238	.206	.180	.157	.138	.122
29	.703	.568	.471	.396	.337	.289	.250	.217	.190	.167	.147	.131
30	.712	.580	.483	.409	.349	.301	.261	.228	.200	.177	.157	.139
40	.779	.668	.583	.513	.455	.406	.364	.327	.295	.267	.243	.221
60	.849	.767	.700	.643	.592	.547	.507	.471	.438	.409	.382	.357
80	.885	.821	.766	.718	.675	.636	.600	.567	.536	.508	.482	.457
100	.908	.854	.809	.768	.730	.696	.664	.634	.606	.580	.555	.532
120	.923	.877	.838	.802	.770	.739	.711	.684	.658	.634	.611	.590
140	.934	.894	.860	.828	.799	.772	.746	.721	.698	.676	.655	.635
170	.945	.912	.883	.856	.831	.808	.785	.764	.743	.724	.705	.687
200	.953	.925	.900	.876	.855	.834	.814	.795	.777	.759	.742	.726
240	.961	.937	.916	.896	.877	.859	.842	.826	.810	.795	.780	.765
320	.971	.952	.936	.921	.907	.893	.879	.866	.854	.841	.829	.818
440	.979	.965	.953	.942	.931	.921	.911	.901	.891	.882	.872	.863
600	.984	.974	.966	.957	.949	.941	.934	.926	.919	.912	.905	.898
800	.988	.981	.974	.968	.961	.956	.950	.944	.938	.933	.927	.922
1000	.991	.985	.979	.974	.969	.964	.960	.955	.950	.946	.941	.937

[a] Multiply entry by 10^{-3}.

(continued)

	ν_H											
ν_E	1	2	3	4	5	6	7	8	9	10	11	12
						$p = 5$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
4	.000	.000	.000	.000	.001^a	.001^a	.001^a	.001^a	.001^a	.001^a	.001^a	.001^a
5	1.60^a	.291^a	.105^a	.052^a	.031^a	.021^a	.015^a	.012^a	.010^a	.008^a	.007^a	.007^a
6	.021	4.39^a	1.48^a	.647^a	.335^a	.197^a	.126^a	.087^a	.064^a	.049^a	.039^a	.032^a
7	.063	.017	6.36^a	2.90^a	1.51^a	.872^a	.544^a	.361^a	.253^a	.185^a	.141^a	.110^a
8	.114	.037	.016	7.74^a	4.21^a	2.48^a	1.56^a	1.03^a	.716^a	.516^a	.385^a	.296^a
9	.165	.063	.029	.015	8.79^a	5.35^a	3.43^a	2.30^a	1.61^a	1.16^a	.861^a	.657^a
10	.215	.092	.046	.026	.015	9.64^a	6.34^a	4.34^a	3.06^a	2.22^a	1.66^a	1.27^a
11	.261	.122	.066	.038	.024	.015	.010	7.22^a	5.17^a	3.80^a	2.86^a	2.19^a
12	.303	.153	.086	.053	.034	.022	.015	.011	7.99^a	5.95^a	4.51^a	3.49^a
13	.341	.183	.108	.068	.045	.031	.022	.016	.012	8.68^a	6.66^a	5.19^a
14	.376	.212	.130	.085	.057	.040	.029	.021	.016	.012	9.31^a	7.32^a
15	.407	.239	.152	.102	.070	.050	.037	.027	.021	.016	.012	9.88^a
16	.436	.266	.174	.119	.084	.061	.045	.034	.026	.020	.016	.013
17	.462	.291	.195	.136	.098	.072	.054	.042	.032	.025	.020	.016
18	.486	.315	.216	.154	.113	.084	.064	.050	.039	.031	.025	.020
19	.508	.337	.236	.171	.127	.096	.074	.058	.046	.037	.030	.024
20	.529	.359	.256	.188	.142	.109	.085	.067	.053	.043	.035	.029
21	.548	.379	.275	.205	.156	.121	.095	.076	.061	.050	.041	.034
22	.565	.398	.293	.221	.171	.134	.106	.085	.069	.057	.047	.039
23	.581	.416	.310	.237	.185	.146	.117	.095	.077	.064	.053	.044
24	.596	.433	.327	.253	.199	.159	.128	.104	.086	.071	.060	.050
25	.610	.449	.343	.268	.213	.171	.139	.114	.094	.079	.066	.056
26	.623	.465	.359	.283	.226	.183	.150	.124	.103	.087	.073	.062
27	.635	.479	.374	.297	.239	.195	.161	.134	.112	.094	.080	.068
28	.647	.493	.388	.311	.252	.207	.172	.143	.121	.102	.087	.075
29	.658	.506	.401	.324	.265	.219	.182	.153	.130	.110	.094	.081
30	.668	.519	.415	.337	.277	.230	.193	.163	.138	.118	.102	.088
40	.744	.617	.522	.446	.384	.333	.291	.255	.224	.198	.176	.156
60	.825	.729	.652	.587	.531	.482	.438	.400	.366	.336	.308	.284
80	.867	.791	.727	.672	.623	.578	.538	.502	.469	.438	.410	.385
100	.893	.830	.776	.728	.685	.645	.609	.576	.544	.516	.489	.464
120	.910	.856	.810	.768	.730	.694	.661	.631	.602	.575	.549	.525
140	.923	.876	.835	.798	.763	.731	.701	.673	.647	.621	.598	.575
170	.936	.897	.862	.830	.801	.773	.747	.722	.698	.675	.654	.633
200	.945	.912	.882	.854	.828	.803	.780	.758	.736	.716	.696	.677
240	.954	.926	.900	.877	.855	.833	.813	.793	.775	.757	.739	.722
300	.966	.944	.925	.906	.889	.872	.856	.841	.825	.811	.797	.783
440	.975	.959	.945	.931	.918	.905	.893	.881	.870	.858	.847	.836
600	.982	.970	.959	.949	.939	.930	.920	.911	.903	.894	.885	.877
800	.986	.977	.969	.961	.954	.947	.940	.933	.926	.919	.913	.906
1000	.989	.982	.975	.969	.963	.957	.951	.946	.940	.935	.929	.924

^{a} Multiply entry by 10^{-3}.

(continued)

	ν_H											
ν_E	1	2	3	4	5	6	7	8	9	10	11	12
						$p = 6$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
4	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
5	.007[a]	.002[a]	.001[a]	.001[a]	.001[a]	.000	.000	.000	.000	.000	.000	.000
6	2.04[a]	.315[a]	.095[a]	.040[a]	.021[a]	.012[a]	.008[a]	.006[a]	.004[a]	.003[a]	.003[a]	.002[a]
7	.019	3.48[a]	1.05[a]	.416[a]	.197[a]	.106[a]	.063[a]	.040[a]	.027[a]	.020[a]	.015[a]	.011[a]
8	.054	.013	4.37[a]	1.82[a]	.872[a]	.465[a]	.270[a]	.168[a]	.111[a]	.076[a]	.055[a]	.041[a]
9	.098	.029	.011	4.94[a]	2.48[a]	1.36[a]	.798[a]	.497[a]	.325[a]	.222[a]	.157[a]	.115[a]
10	.144	.050	.021	.010	5.35[a]	3.04[a]	1.83[a]	1.16[a]	.762[a]	.521[a]	.369[a]	.269[a]
11	.189	.074	.034	.017	9.64[a]	5.67[a]	3.51[a]	2.26[a]	1.51[a]	1.05[a]	.744[a]	.543[a]
12	.232	.099	.049	.027	.015	9.35[a]	5.94[a]	3.92[a]	2.66[a]	1.86[a]	1.34[a]	.983[a]
13	.271	.126	.066	.037	.022	.014	9.17[a]	6.17[a]	4.27[a]	3.03[a]	2.20[a]	1.63[a]
14	.308	.152	.084	.049	.031	.020	.013	9.07[a]	6.38[a]	4.59[a]	3.37[a]	2.52[a]
15	.341	.179	.103	.063	.040	.026	.018	.013	9.00[a]	6.57[a]	4.88[a]	3.68[a]
16	.372	.204	.122	.077	.050	.034	.024	.017	.012	8.97[a]	6.74[a]	5.14[a]
17	.400	.229	.141	.091	.061	.042	.030	.021	.016	.012	8.97[a]	6.90[a]
18	.426	.252	.160	.106	.072	.051	.037	.027	.020	.015	.012	8.97[a]
19	.450	.275	.179	.121	.084	.060	.044	.033	.025	.019	.015	.011
20	.473	.296	.197	.136	.096	.070	.052	.039	.030	.023	.018	.014
21	.493	.317	.215	.151	.109	.080	.060	.045	.035	.027	.021	.017
22	.512	.337	.233	.166	.121	.090	.068	.052	.041	.032	.025	.020
23	.530	.355	.250	.181	.134	.101	.077	.060	.047	.037	.030	.024
24	.546	.373	.266	.195	.146	.111	.086	.067	.053	.042	.034	.028
25	.562	.390	.282	.210	.159	.122	.095	.075	.060	.048	.039	.032
26	.576	.406	.298	.224	.171	.133	.104	.083	.066	.054	.044	.036
27	.590	.422	.313	.237	.183	.143	.113	.091	.073	.060	.049	.040
28	.603	.436	.327	.251	.195	.154	.123	.099	.080	.066	.054	.045
29	.615	.450	.341	.264	.207	.165	.132	.107	.088	.072	.060	.050
30	.626	.464	.355	.277	.219	.175	.142	.116	.095	.079	.066	.055
40	.711	.570	.467	.387	.324	.273	.232	.198	.170	.147	.127	.110
60	.802	.693	.608	.536	.476	.424	.379	.340	.305	.275	.249	.225
80	.849	.762	.690	.629	.574	.526	.483	.445	.410	.378	.350	.324
100	.878	.806	.745	.691	.642	.599	.559	.523	.489	.458	.430	.404
120	.898	.836	.783	.735	.692	.652	.616	.582	.551	.521	.494	.468
140	.912	.858	.811	.769	.730	.694	.660	.629	.599	.572	.546	.521
170	.927	.882	.842	.806	.772	.740	.710	.682	.656	.630	.607	.584
200	.938	.899	.864	.832	.803	.774	.748	.722	.698	.675	.653	.632
240	.948	.915	.886	.858	.833	.808	.785	.763	.741	.721	.701	.682
320	.961	.936	.913	.892	.872	.852	.834	.816	.799	.782	.766	.750
440	.972	.953	.936	.920	.905	.890	.876	.862	.849	.836	.823	.811
600	.979	.965	.953	.941	.930	.918	.908	.897	.887	.877	.867	.857
800	.984	.974	.964	.955	.947	.938	.930	.922	.914	.906	.898	.891
1000	.987	.979	.971	.964	.957	.950	.944	.937	.930	.924	.918	.912

[a] Multiply entry by 10^{-3}.

(continued)

	ν_H											
ν_E	1	2	3	4	5	6	7	8	9	10	11	12
						$p = 7$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
4	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
5	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
6	.043[a]	.006[a]	.002[a]	.001[a]	.001[a]	.000	.000	.000	.000	.000	.000	.000
7	2.62[a]	.350[a]	.091[a]	.033[a]	.015[a]	.008[a]	.005[a]	.003[a]	.002[a]	.002[a]	.001[a]	.001[a]
8	.018	2.95[a]	.809[a]	.292[a]	.126[a]	.063[a]	.034[a]	.020[a]	.013[a]	.009[a]	.006[a]	.005[a]
9	.048	.010	3.20[a]	1.22[a]	.543[a]	.270[a]	.147[a]	.086[a]	.053[a]	.035[a]	.024[a]	.017[a]
10	.087	.023	8.07[a]	3.34[a]	1.56[a]	.798[a]	.440[a]	.259[a]	.160[a]	.104[a]	.070[a]	.049[a]
11	.128	.040	.016	6.97[a]	3.43[a]	1.83[a]	1.04[a]	.619[a]	.387[a]	.252[a]	.170[a]	.119[a]
12	.170	.060	.026	.012	6.34[a]	3.51[a]	2.05[a]	1.25[a]	.796[a]	.525[a]	.357[a]	.249[a]
13	.209	.083	.038	.019	.010	5.94[a]	3.57[a]	2.23[a]	1.45[a]	.967[a]	.665[a]	.468[a]
14	.246	.106	.052	.027	.015	9.17[a]	5.67[a]	3.63[a]	2.40[a]	1.62[a]	1.13[a]	.804[a]
15	.281	.129	.067	.037	.022	.013	8.37[a]	5.48[a]	3.68[a]	2.54[a]	1.79[a]	1.28[a]
16	.313	.153	.083	.047	.029	.018	.012	7.80[a]	5.34[a]	3.73[a]	2.66[a]	1.94[a]
17	.343	.176	.099	.059	.037	.024	.016	.011	7.38[a]	5.24[a]	3.78[a]	2.78[a]
18	.370	.199	.116	.071	.045	.030	.020	.014	9.81[a]	7.06[a]	5.16[a]	3.83[a]
19	.396	.221	.133	.083	.054	.037	.025	.018	.013	9.20[a]	6.80[a]	5.10[a]
20	.420	.242	.149	.096	.064	.044	.031	.022	.016	.012	8.72[a]	6.60[a]
21	.442	.263	.166	.109	.074	.052	.037	.026	.019	.014	.011	8.34[a]
22	.462	.283	.183	.123	.085	.060	.043	.031	.023	.018	.013	.010
23	.482	.301	.199	.136	.095	.068	.050	.037	.028	.021	.016	.013
24	.499	.320	.215	.149	.106	.077	.057	.042	.032	.025	.019	.015
25	.516	.337	.230	.162	.117	.086	.064	.048	.037	.029	.022	.018
26	.532	.354	.246	.175	.128	.095	.071	.055	.042	.033	.026	.020
27	.547	.370	.260	.188	.139	.104	.079	.061	.047	.037	.029	.024
28	.561	.385	.275	.201	.150	.113	.087	.068	.053	.042	.033	.027
29	.574	.399	.289	.214	.161	.123	.095	.074	.059	.047	.037	.030
30	.586	.413	.302	.226	.172	.132	.103	.081	.064	.052	.042	.034
40	.679	.526	.417	.335	.273	.224	.185	.154	.128	.108	.091	.077
60	.779	.660	.566	.490	.426	.373	.327	.288	.254	.225	.200	.178
80	.832	.735	.656	.588	.530	.479	.434	.394	.358	.326	.298	.272
100	.864	.783	.715	.656	.603	.556	.513	.475	.439	.408	.378	.352
120	.886	.817	.757	.704	.657	.613	.574	.537	.504	.473	.444	.418
140	.902	.841	.788	.741	.698	.658	.621	.587	.556	.526	.498	.472
170	.919	.868	.823	.782	.744	.709	.676	.645	.616	.589	.563	.539
200	.931	.887	.848	.812	.778	.747	.717	.689	.662	.637	.613	.590
240	.942	.905	.871	.841	.812	.784	.758	.733	.709	.687	.665	.644
320	.957	.928	.902	.878	.855	.833	.812	.792	.773	.754	.736	.719
440	.968	.947	.928	.910	.893	.876	.860	.844	.829	.814	.800	.786
600	.977	.961	.947	.933	.920	.908	.895	.883	.872	.860	.849	.838
800	.982	.971	.960	.950	.940	.930	.920	.911	.902	.893	.884	.876
1000	.986	.977	.968	.959	.951	.943	.936	.928	.921	.914	.906	.899

[a] Multiply entry by 10^{-3}.

(continued)

	ν_H											
ν_E	1	2	3	4	5	6	7	8	9	10	11	12
						$p = 8$						
1	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
4	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
5	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
6	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
7	.138[a]	.015[a]	.004[a]	.001[a]	.001[a]	.000	.000	.000	.000	.000	.000	.000
8	3.30[a]	.393[a]	.090[a]	.029[a]	.012[a]	.006[a]	.003[a]	.002[a]	.001[a]	.001[a]	.001[a]	.000
9	.017	2.63[a]	.659[a]	.218[a]	.087[a]	.040[a]	.020[a]	.011[a]	.007[a]	.004[a]	.003[a]	.002[a]
10	.044	8.63[a]	2.46[a]	.872[a]	.361[a]	.168[a]	.086[a]	.047[a]	.028[a]	.017[a]	.011[a]	.008[a]
11	.078	.019	6.15[a]	2.36[a]	1.03[a]	.497[a]	.259[a]	.144[a]	.085[a]	.052[a]	.034[a]	.023[a]
12	.116	.033	.012	4.99[a]	2.30[a]	1.16[a]	.619[a]	.351[a]	.209[a]	.130[a]	.084[a]	.056[a]
13	.154	.051	.020	8.91[a]	4.34[a]	2.26[a]	1.25[a]	.727[a]	.441[a]	.278[a]	.181[a]	.122[a]
14	.190	.070	.030	.014	7.22[a]	3.92[a]	2.23[a]	1.33[a]	.824[a]	.527[a]	.347[a]	.235[a]
15	.225	.090	.041	.021	.011	6.17[a]	3.63[a]	2.22[a]	1.40[a]	.910[a]	.608[a]	.416[a]
16	.258	.111	.054	.028	.016	9.06[a]	5.48[a]	3.42[a]	2.20[a]	1.46[a]	.987[a]	.683[a]
17	.289	.133	.067	.037	.021	.013	7.80[a]	4.98[a]	3.27[a]	2.20[a]	1.51[a]	1.06[a]
18	.318	.154	.082	.046	.027	.017	.011	6.92[a]	4.62[a]	3.15[a]	2.19[a]	1.56[a]
19	.345	.175	.096	.056	.034	.021	.014	9.23[a]	6.26[a]	4.34[a]	3.06[a]	2.19[a]
20	.370	.195	.111	.067	.042	.027	.018	.012	8.22[a]	5.77[a]	4.12[a]	2.99[a]
21	.393	.215	.127	.078	.050	.033	.022	.015	.010	7.46[a]	5.39[a]	3.95[a]
22	.415	.235	.142	.089	.058	.039	.026	.018	.013	9.40[a]	6.86[a]	5.08[a]
23	.436	.254	.157	.101	.067	.045	.031	.022	.016	.012	8.56[a]	6.39[a]
24	.455	.272	.172	.113	.076	.052	.037	.026	.019	.014	.010	7.88[a]
25	.473	.289	.187	.124	.085	.060	.042	.031	.023	.017	.013	9.56[a]
26	.490	.306	.201	.136	.095	.067	.048	.035	.026	.020	.015	.011
27	.505	.322	.215	.148	.104	.075	.055	.040	.030	.023	.017	.013
28	.520	.338	.229	.160	.114	.083	.061	.045	.034	.026	.020	.016
29	.534	.353	.243	.172	.124	.091	.068	.051	.039	.030	.023	.018
30	.548	.367	.256	.183	.134	.099	.074	.056	.043	.034	.026	.021
40	.649	.485	.372	.290	.229	.182	.146	.118	.096	.079	.065	.054
60	.758	.627	.527	.447	.381	.327	.282	.244	.212	.184	.161	.141
80	.815	.709	.623	.551	.489	.435	.389	.348	.313	.281	.253	.229
100	.851	.761	.687	.622	.566	.516	.471	.431	.395	.362	.333	.306
120	.875	.798	.732	.675	.623	.577	.535	.496	.461	.429	.399	.372
140	.892	.825	.767	.715	.667	.625	.585	.549	.515	.484	.455	.428
170	.911	.854	.804	.759	.717	.679	.644	.610	.579	.550	.523	.497
200	.924	.875	.831	.791	.755	.720	.688	.657	.629	.602	.576	.551
240	.936	.895	.858	.823	.791	.761	.732	.705	.679	.655	.631	.609
320	.952	.920	.891	.865	.839	.815	.792	.770	.748	.728	.708	.689
440	.965	.942	.920	.900	.880	.862	.844	.827	.810	.794	.778	.762
600	.974	.957	.941	.926	.911	.897	.883	.870	.857	.844	.831	.819
800	.981	.968	.955	.944	.933	.922	.911	.901	.890	.880	.871	.861
1000	.985	.974	.964	.955	.946	.937	.928	.920	.911	.903	.895	.887

[a] Multiply entry by 10^{-3}.

توجه داشته باشید که می توان با تبدیلات ذیل این آماره را به آماره F تبدیل کرد:

جدول b) تبدیلات لاندای ویلکس به دقیق ترین دم بالای آزمون F

Parameters p, ν_H	Statistic Having F-Distribution	Degrees of Freedom
Any $p, \nu_H = 1$	$\frac{1-\Lambda}{\Lambda}\frac{\nu_E - p + 1}{p}$	$p, \nu_E - p + 1$
Any $p, \nu_H = 2$	$\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\frac{\nu_E - p + 1}{p}$	$2p, 2(\nu_E - p + 1)$
$p = 1$, any ν_H	$\frac{1-\Lambda}{\Lambda}\frac{\nu_E}{\nu_H}$	ν_H, ν_E
$p = 2$, any ν_H	$\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\frac{\nu_E - 1}{\nu_H}$	$2\nu_H, 2(\nu_E - 1)$

برای مقادیری از p و ν_H که در جدول بالا موجود نیست می توان از تقریب زیر استفاده کرد :

$$F = \frac{1-\Lambda^{1/t}}{\Lambda^{1/t}}\frac{\mathrm{df}_2}{\mathrm{df}_1},$$

که در آن مقادیر مربوطه بصورت زیر است :

$$\mathrm{df}_1 = p\nu_H, \qquad \mathrm{df}_2 = wt - \frac{1}{2}(p\nu_H - 2),$$

$$w = \nu_E + \nu_H - \frac{1}{2}(p + \nu_H + 1), \qquad t = \sqrt{\frac{p^2\nu_H^2 - 4}{p^2 + \nu_H^2 - 5}}.$$

همانطور که از روابط بالا مشخص است می توان گفت آماره لاندای ویلکس و آمارهF معکوس یکدیگرند و در تمامی آزمون هایی که ما در این پروژه انجام دادیم می توان مقدار آماره را تعویض کرد و بنابراین جهت های ناحیه بحرانی نیز عوض می شوند.

منابع

منابع فارسی :

۱. تحلیل آماری چند متغیری کاربردی،ریچارد آ.جانسون،دین دبلیو.ویچرن ترجمـه دکتـر حسینعلی نیرومند

منابع لاتین :

1. Alvin C. Rencher ,Methods Of Multivariate Analysis, second edition , WHILEY SERIES IN PROBABILITY AND STATISTICS
2. Ravindra khatteree & Dayanand N.Nike, Applied Multivariate Statistics With SAS software ,second edition , SAS® and all other SAS
3. Neil H. Timm,Applied Multivariate Analysis, SPRINGER TEXT IN STATTISTICS
4. Goldberger, A.S. (1991). *A course in econometrics.* Cambridge, MA:Harvard University Press.
5. Neter, J., Kutner, M.H., Nachtsheim, C.J. andWasserman,W. (1996). *Applied linear statistical models, regression, analysis of variance and experimental designs*. Fourth Edition. Homewood, IL: Richard D. Irwin, Inc.
6. Timm, N.H. and Mieczkowski, T.A. (1997). *Univariate and multivariate general linear models: Theory and applications using SAS software*. Cary, NC: SAS Institute Inc