

سری سوال: یک ۱

زمان آزمون (دقیقه): تستی: ۶۰ تشریحی: ۶۰

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

دروس: رگرسیون ۲

رشته تحصیلی/ گد درس: آمار و کاربردها ۱۱۱۷۱۸۰

۱- باقیمانده های استاندارد شده کدامند؟

$r_i \cdot .^4$

$\frac{e_i}{\sqrt{MSE}} \cdot .^3$

$\frac{e_i}{\delta} \cdot .^2$

$e_i \cdot .^1$

۲- برای کنترل فرض نرمال بودن باقیمانده استفاده از نمودار است.

۱. احتمال نرمال باقیمانده ها

۱. باقیمانده ها نسبت به مقادیر پیش‌بینی

۴. باقیمانده ها در مقابل X_i ها

۳. باقیمانده ها در مقابل مقادیر متناظر متغیر رگرسیون

۳- اگر نمودار باقیمانده ها در مقابل \hat{Y}_i ها به صورت قیفی شکل باز شده به طرف خارج باشد حکایت از آن دارد که۲. واریانس تابع نزولی از Y است.۱. واریانس تابع صعودی از Y است.

۴. واریانس خطاهای ثابت نیست.

۳. اگر Y کاهش یابد (\mathcal{E}) V افزایش یابد.

۴- روش حداقل مربعات موزون در چه موقعی کاربرد دارد؟

۱. مدلهای رگرسیونی با واریانس خطای ثابت

۱. مدلهای رگرسیونی داده های دورافتاده

۴. مدلهای رگرسیونی با واریانس خطای غیر ثابت

۳. مدلهای رگرسیونی با واریانس خطای نامتغیر

۵- برآورد حداقل مربعات خطای برای برآورد پارامترهای مدل رگرسیونی کدام است؟

$\hat{\beta} = (X X')^{-1} \cdot X' Y \cdot .^2$

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

$\hat{\beta} = X' (X X')^{-1} \cdot X' Y \cdot .^4$

$\hat{\beta} = X (X' X)^{-1} \cdot X' Y \cdot .^3$

۶- اگر ماتریس (X') موجود نباشد یعنی

۱. داده های دورافتاده وجود دارد.

۲. متغیرهای رگرسیونی به طور خطی از هم مستقل‌اند.

۳. متغیرهای رگرسیونی به طور خطی از هم مستقل نمی‌باشند.

۴. داده های با نفوذ وجود دارد.

سری سوال: ۱ یک

زمان آزمون (دقیقه): تستی: ۶۰ تشریحی: ۵

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

دوس: رگرسیون ۲

روش تحلیلی/گد درس: آمار و کاربردها ۱۱۱۷۱۸۰

۷- طبق قضیه گوس- مارکف داریم که

۱. واریانس خطاهای معمولاً ناریب است.

۲. برآوردگر درست نمایی ماکزیمم بهترین برآوردگر برای پارامتر β است.۳. برآوردگر حداقل مربعات $\hat{\beta}$ بهترین برآوردگر خطی پارامترهای β است.۴. برآوردگر حداقل مربعات $\hat{\beta}$ بهترین برآوردگر خطی ناریب β است.۸- برای یک مدل رگرسیونی با مشخصات زیر برآورد σ کدام است؟

$$SSE = 3VQ, P = 2, n = 2Q$$

۱۷/۰۵ .۴

۱۶/۳ .۳

۱۵ .۲

۷/۵ .۱

۹- اگر برآورد نقطه‌ای پارامتر β از یک مدل رگرسیون خطی چندگانه برابر با $1/62$ و عضو قطعی متناظر از $(X'X)^{-1}$ برای۱۰- $C_{11} = \hat{\sigma}^2$ باشد. حد بالای یک فاصله اطمینان ۹۵٪ برای پارامتر β کدام است. (عدد جدول

مورد نیاز برابر با ۲۰۷۴ می باشد)

۳/۲۱ .۴

۲/۸۹ .۳

۲/۴۱ .۲

۳/۷ .۱

۱۰- عبارت نادرست کدام است؟۱. معیار R^2 به عنوان یک معیار کاهش در تغییر پذیری Y که از به کارگیری متغیرهای رگرسیون i ها در مدل بدست می آید در نظر گرفته می شود.۲. اضافه کردن یک متغیر به مدل R^2 را افزایش می دهد.۳. R^2 مربع همبستگی بین Y , \hat{Y} است.۴. یک مقدار R^2 بزرگ لزوماً یک رگرسیون خوب را نتیجه می دهد.۱۱- رابطه بین R^2 , \bar{R}^2 کدام است؟

$$\bar{R}^2 = 1 + \frac{n-1}{n-p} (R^2 - 1) \quad .\cdot ۱$$

$$\bar{R}^2 = 1 - \frac{n-p}{n-1} (R^2 - 1) \quad .\cdot ۱$$

$$R^2 = 1 - \frac{n-p}{n-1} \bar{R}^2 \quad .\cdot ۴$$

$$R^2 = 1 - \frac{n-1}{n-p} \bar{R}^2 \quad .\cdot ۳$$

سری سوال: ۱ یک

زمان آزمون (دقیقه): تستی: ۶۰ تشریحی: ۵

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

دوس: رگرسیون ۲

روش تحلیلی/گد درس: آمار و کاربردها ۱۱۱۷۱۸۰

۱۲- جهت شناسایی هم خطی از کدام معیار زیر استفاده می کنیم؟

۲. آماره PRESS

۱. فاصله کوک

۴. ضریب تعیین تعدیل شد

۳. عامل تورم واریانس

۱۳- در چه موقعی از توابع اسپلین استفاده میشود؟

۱. چند جمله ای از مرتبه پایین بر ازش ضعیفی داشته باشد.

۲. گرایش پیدا نکردن مجموع مربعات باقیمانده ها به مقدار ثابت

۳. هرگاه نمودارهای باقیمانده ها نمایش دهنده ساختمان توضیح داده شده بوسیله مدل باشند.

۴. همه موارد

۱۴- یک متغیر کیفی با a سطح بوسیلهمتغیر نشانگر ارایه می شود که هریک مقادیررا می گیرند.۲. $(a - 1)$ ، کیفی۱. a ، رسته ای۴. a ، صفر و یک۳. $(a - 1)$ ، صفر و یک۱۵- مشاهده i ام یک نقطه با تاثیر بالا تلقی می شود هرگاه h_{ii} (قطر اعضای ماتریس H)باشد وقتی که p تعداد متغیرها و n تعداد مشاهدات است.۴. $\frac{\sum p}{n}$ کمتر از۳. $\frac{\sum p}{n}$ بیشتر از۲. $\frac{p}{n}$ کمتر از۱. $\frac{p}{n}$ بیشتر از۱۶- در مورد ماتریس $H = X(X'X)^{-1}X'$ برای مدل رگرسیونی $Y = X\beta + e$ و باقیمانده های e کدام عبارت صحیح نیست؟۴. $\in = (I - H)e$ ۳. $e = (I - H)\in$ ۲. $H = H^2$ ۱. $H = H'$ ۱۷- کدام عبارت در مورد جمله خطای در مدل رگرسیونی صادق نیست؟

۲. خطای دارای میانگین صفر است.

۱. خطای دارای میانگین صفر است.

۴. خطای ناهمبسته است.

۳. خطای دارای همبستگی است.

سری سوال: ۱ یک

زمان آزمون (دقیقه): تستی: ۶۰ تشریحی: ۶۰

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

دوس: رگرسیون ۲

روش تحلیلی/گد درس: آمار و کاربردها ۱۱۱۷۱۸۰

۱۸- ملاک ارزیابی مدل رگرسیونی C_p بدین صورت است که

۱. معادلات رگرسیونی با اربیبی زیاد مقادیری از $C_p = p$ خواهند داشت که بالاتر از خط $C_p = p$ است.
۲. معادلات رگرسیونی با اربیبی کم مقادیری از $C_p = p$ خواهند داشت که نزدیک خط $C_p = p$ است.
۳. معادلات رگرسیونی با اربیبی زیاد مقادیری از $C_p = p$ خواهند داشت که زیر خط $C_p = p$ است.
۴. معادلات رگرسیونی با اربیبی متوسط مقادیری از $C_p = p$ خواهند داشت که نزدیک خط $C_p = p$ است.

۱۹- کدامیک از این روشها برای انتخاب بهترین مدل رگرسیونی نمی باشد؟

۱. روش رگرسیون قدم به قدم
۲. روش رگرسیون پیش رو
۳. روش رگرسیون حذف پس رو

۲۰- کاربرد روش رگرسیون مضرس عبارت است از

۱. کاهش اثرات داده های دور افتاده
۲. باقی براورد های دقیق برای پارامترها
۳. بدست آوردن برآوردگرهای اریب ضرایب رگرسیون

سوالات تشریحی

۱۴ نمره

۱- مدل های غیر خطی زیر را با تبدیلات مناسب خطی کنید.

$$\text{الف-} Y = \beta_0 e^{\beta_1 X} \varepsilon$$

$$\text{ب-} Y = \frac{X}{\beta_0 + \beta_1 x}$$

$$\text{ج-} Y = \beta_0 + \beta_1 x$$

سری سوال: ۱ یک

زمان آزمون (دقیقه): تستی: ۶۰ تشریحی: ۶۰

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

دوس: رگرسیون ۲

روش تحلیلی/گد درس: آمار و کاربردها ۱۱۱۷۱۸۰

۱.۴ نمره

 $E \sim N(0, \sigma^2) = \beta_0 x_1 + \beta_1 x_2 + \beta_2 x_3 + \epsilon$ مدل خطی زیر را در نظر بگیرید

است یافته های یک نمونه تصادفی در جدول زیر داده شده است.

| Y | X_1 | X_2 | X_3 |
|----|-------|-------|-------|
| ۰ | ۱ | -۱ | ۰ |
| ۱ | ۰ | ۰ | ۱ |
| -۱ | ۱ | ۰ | ۰ |
| ۲ | ۱ | ۱ | ۰ |

الف- برآورد پارامترها را بدست آورید برآورد و همچنین برآورد $\hat{\beta}$ را بیابید؟

ب- ماتریس H (هـ) یا ماتریس برازش را تشکیل دهید.

ج- باقیمانده ها را بدست آورید. برآورد $\hat{\sigma}^2$ را بدست آورید.

۱.۴ نمره

۳- ضریب تعیین چند گانه و ضریب تعیین تعدیل شده را تعریف کرده و تفاوت آنها و نحوه کاربردشان را توضیح دهید.

۱.۴ نمره

۴- با مراجعه به مسئله (۲) مطلوب است محاسبه باقیمانده های استیودنت شده (r_i) و آماره PRESS ؟

۱.۴ نمره

۵- کاربرد باقیمانده ها (e_i) در مدل رگرسیون چند گانه جهت قضاوت در مورد مناسب بودن مدل را توضیح داده و در مورد نمودار باقیمانده های جزیی و کاربرد آن توضیح دهید.