

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: یک ۱

درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها ۱۱۷۱۸۰

۱- باقیمانده های استاندارد شده کدامند؟

$$e_i \quad .1 \quad \frac{e_i}{\delta} \quad .2 \quad \frac{e_i}{\sqrt{MSE}} \quad .3 \quad r_i \quad .4$$

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

۱. باقیمانده ها نسبت به مقادیر پیش بینی
۲. احتمال نرمال باقیمانده ها
۳. باقیمانده ها در مقابل مقادیر متناظر متغیر رگرسیون
۴. باقیمانده ها در مقابل X_i ها

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 که

۱. واریانس تابع صعودی از Y است.
۲. واریانس تابع نزولی از Y است.
۳. اگر Y کاهش یابد $V(\epsilon)$ افزایش یابد.
۴. واریانس خطاها ثابت نیست.

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

۱. مدل های رگرسیونی با واریانس خطای ثابت
۲. برای شناسایی داده های دورافتاده
۳. مدل های رگرسیونی با واریانس خطای نامتغیر
۴. مدل های رگرسیونی با واریانس خطای غیر ثابت

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

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

$$\hat{\beta} = X(X'X)^{-1} X'Y \quad .3 \quad \hat{\beta} = X'(X'X)^{-1} X'Y \quad .4$$

۶- اگر ماتریس $(X'X)$ موجود نباشد یعنی

۱. داده های دورافتاده وجود دارد.
۲. متغیرهای رگرسیونی به طور خطی از هم مستقل اند.
۳. متغیرهای رگرسیونی به طور خطی از هم مستقل نمی باشند.
۴. داده های با نفوذ وجود دارد.

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها ۱۱۷۱۸۰

۷- طبق قضیه گوس - مارکف داریم که

۱. واریانس خطاها 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 $\hat{\beta}$ کدام است؟

$$SSE = 275, P = 2, n = 25$$

۱. ۷/۵ ۲. ۱۵ ۳. ۱۶/۳ ۴. ۱۷/۰۵

۹- اگر برآورد نقطه ای پارامتر β_1 از یک مدل رگرسیون خطی چندگانه برابر با ۱/۶۲ و عضو قطری متناظر از $(X'X)^{-1}$ برای

β_1 برابر با ۰/۱، $C_{11} = ۱۰$ ، $\sigma^2 = ۱۰$ باشد. حد بالای یک فاصله اطمینان ۹۵٪ برای پارامتر β_1 کدام است. (عدد جدول مورد نیاز برابر با ۲/۰۷۴ می باشد)

۱. ۳/۷ ۲. ۲/۴۱ ۳. ۲/۸۹ ۴. ۳/۲۱

۱۰- عبارت نادرست کدام است؟

۱. معیار R^2 به عنوان یک معیار کاهش در تغییر پذیری Y که از به کارگیری متغیرهای رگرسیون X_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 \hat{Y} ، 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؟

$$\begin{aligned} \bar{R}^2 &= 1 - \frac{n-1}{n-p} (R^2 - 1) & \bar{R}^2 &= 1 - \frac{n-p}{n-1} (R^2 - 1) \\ R^2 &= 1 - \frac{n-p}{n-1} \bar{R}^2 & R^2 &= 1 - \frac{n-1}{n-p} \bar{R}^2 \end{aligned}$$

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها ۱۱۷۱۸۰

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

۱. فاصله کوک
۲. آماره PRESS
۳. عامل تورم واریانس
۴. ضریب تعیین تعدیل شد

۱۳- در چه مواقعی از توابع اسپلین استفاده میشود؟

۱. چند جمله ای از مرتبه پایین بر ارزش 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.

۱. a ، رسته ای
۲. $(a - 1)$ ، کیفی
۳. $(a - 1)$ ، صفر و یک
۴. a ، صفر و یک

۱۵- 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.

۱. $\frac{p}{n}$ بیشتر از
۲. $\frac{p}{n}$ کمتر از
۳. $\frac{\sum p}{n}$ بیشتر از
۴. $\frac{\sum p}{n}$ کمتر از

۱۶- در مورد ماتریس $X'X$ $H = X(X'X)^{-1}X'$ برای مدل رگرسیونی $Y = X\beta + \epsilon$ و باقیمانده های e کدام عبارت صحیح نیست؟

۱. $H = H'$
۲. $H = H^2$
۳. $e = (I - H)\epsilon$
۴. $\epsilon = (I - H)e$

۱۷- کدام عبارت در مورد جمله خطا در مدل رگرسیونی صادق نیست؟

۱. خطا دارای میانگین صفر است.
۲. خطا دارای واریانس ثابت است.
۳. خطاها دارای همبستگی اند.
۴. خطاها ناهمبسته اند.

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها ۱۱۷۱۸۰

۱۸- ملاک ارزیابی مدل رگرسیونی C_p بدین صورت است که

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

۲۰- کاربرد روش رگرسیون مضرس عبارت است از

۱. کاهش اثرات داده های دور افتاده
۲. بافتی برآوردهای دقیق برای پارامترها
۳. بدست آوردن برآوردهای اریب ضرایب رگرسیون
۴. از بین بردن اثرات همبستگی درون داده ها

سوالات تشریحی

۱- مدلهای غیر خطی زیر را با تبدیلات مناسب خطی کنید.

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

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

ج- $Y = \beta_0 x^{\beta_1}$

۱.۴ نمره

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها ۱۱۷۱۸۰

۱.۴ نمره

۲- مدل خطی زیر را در نظر بگیرید $Y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$ به طوری که $E \sim N(0, \sigma^2)$ 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{Y} را بیابید؟

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

ج- باقیمانده ها را بدست آورید. برآورد $\hat{\epsilon}$ 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 مسأله (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 جزئی و کاربرد آن توضیح دهید.