



تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: یک ۱

عنوان درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها، ریاضیات و کاربردها ۱۱۷۱۸۰

۱- مقدار  $\text{var}(\hat{y}_i) - \text{cov}(y_i - \hat{y}_i)$  کدام گزینه است؟

۱. ۱  $\sigma^2$  .۲

۳. ۰  $\sigma^2 \left( \frac{1}{n} + \frac{x_i - x_p}{S_{XX}} \right)$  .۴

۲- عبارت  $\frac{e_i}{\sqrt{MSE}}$  که در آن  $MSE = \frac{SSE}{n-2}$  کدام گزینه است؟

۱. باقی مانده استاندارد شده  
۲. باقی مانده استیودنت شده  
۳. باقی مانده های حداقل و مربعات  
۴. باقی مانده های مقیاس نشده

۳- در مدل رگرسیون با  $n_i$  تکرار برای هر مقدار  $x_i$  عبارت  $\sum_{i=1}^m n_i (\bar{y}_i - \hat{y}_i)^2$  کدام عبارت را محاسبه می کنند؟

۱. مجموع مربعات خطاها  
۲. مجموع مربعات خطای محض  
۳. مجموع مربعات کمبود برازش  
۴. مجموع مربعات کل

۴- در مدل رگرسیون چندمتغیره اگر  $H = X(X'X^{-1})X'$  باشد بردار خطای  $e$  با کدام گزینه برابر است؟

۱.  $Hy$  .۲  $I - HY$  .۳  $(I - H)y$  .۴  $H - Hy$

۵- اگر  $\hat{\beta}$  برآورد حداقل و مربعات  $\beta$  باشد کدام گزینه نادرست است؟

۱. اگر  $\epsilon_i$  دارای توزیع نرمال باشد آنگاه  $\hat{\beta}$  برآوردگر درستنمایی ماکزیمم هم خواهد بود.  
۲. طبق قضیه گرس - مارکف  $\hat{\beta}$  بهترین برآوردگر خطی ناریب  $\beta$   
۳. اگر  $\epsilon_i$  دارای توزیع نرمال باشد آن گاه  $\hat{\beta}$  بهترین برآوردگر خطی ناریب است.  
۴. طبق قضیه گوس - مارکف برآوردگر درستنمایی ماکزیمم و برآوردگر حداقل مربعات یکی هستند.



تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

عنوان درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها، ریاضیات و کاربردها ۱۱۷۱۸۰

۶- جدول آنالیز واریانس یک مدل رگرسیونی را در نظر بگیرید

منبع تغییرات	مجموع مربعات	درجات آزادی	میانگین مربعات	$F_0$
رگرسیون			۵۰	۲
باقی مانده ها				
کل	۵۲۵			

اگر تعداد متغیرهای مستقل رگرسیونی ۳ باشد مقدار مجموع مربعات باقی مانده ها چقدر است؟

۱۸۴/۵ .۴

۸۵ .۳

۴۲۵ .۲

۲۶۲/۵ .۱

۷- در آماره آزمون برای کمبود برازش  $F_0 = \frac{MS_{LOF}}{MS_{PE}}$  کدام گزاره درست نمی باشد؟

۱. اگر  $F_0 \geq F_{\alpha, m-2, n-m}$  تابع رگرسیون خطی نیست.۲. آماره  $F_0$  دارای توزیع  $F_{m-2, n-m}$  می باشد.۳.  $E(MS_{LOF}) = E(MS_{PE})$ 

۴. همواره تابع رگرسیون خطی نیست.

۸- مدل  $y = \beta_0 e^{\beta_1 x}$  میتواند با تبدیل  $y' = \ln y$  به کدام مدل تبدیل شود؟۱.  $y' = \ln \beta_0 + \beta_1 x$ ۲.  $y' = \beta_0 + \beta_1 x$ ۳.  $y' = \ln \beta_0 + \beta_1 \ln x$ ۴.  $y' = \beta_0 + \beta_1 \ln x$ 

۹- کدام گزینه درباره تبدیلات تثبیت کننده واریانس درست است؟

۱. اگر  $y$  دارای توزیع پواسون باشد می توان  $y' = \sqrt{y}$  را در مقابل  $x$  در نظر گرفت.

۲. برای یک دامنه تغییرات نسبتا باریک نمی توان یک تبدیل ملایم بکار برد.

۳. ثابت بودن واریانس یک ضرورت اساسی در تحلیل رگرسیون نمی باشد.

۴. برای اینکه مدلی به یک مدل مناسب بهبود یابد و تصحیح شود یک تبدیل ضرورت دارد ولی این تبدیل به تثبیت واریانس کمک نمی کند.

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

عنوان درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها، ریاضیات و کاربردها ۱۱۷۱۸۰

۱۰- اگر در مدل  $y = x\beta + \varepsilon$  بردار ضرایب  $\beta$  بصورت  $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$  صورت پذیرد مدل بصورت

$y = x\beta + \varepsilon = x_1\beta_1 + x_p\beta_p + \varepsilon$  تبدیل میشود. با شرط آنکه متغیرهای  $x_1$  بر متغیرهای  $x_p$  عمود هستند  $SSR(\beta_p | \beta_1)$  کدام گزینه است؟

۱.  $SSR(\beta_1)$       ۲.  $SSR(\beta_p) - SSR(\beta_1)$

۳.  $SSR(\beta_p)$       ۴.  $SSR(\beta_1) - SSR(\beta_p)$

۱۱- مدل  $y = \beta_0 + \beta_1 x_1 + \beta_p x_p + \beta_{11} x_1^2 + \beta_{pp} x_p^2 + \beta_{1p} x_1 x_p + \varepsilon$  را در سطر نظر بگیرید این مدل را چه می نامند؟

۱. چندجمله ای یک متغیره درجه ۲      ۲. چندجمله ای دو متغیره درجه یک

۳. چندجمله ای دو متغیره درجه ۲      ۴. چندجمله ای یک متغیره درجه یک با اثر متقابل

۱۲- در هنگام برازش یک چندجمله ای یک متغیره وقتی که  $n$  مشاهده داریم حداکثر درجه چندجمله ای چند باشد تا پیش بینی به خوبی صورت پذیرد؟

۱. کمتر از یک      ۲. کوچکتر یا مساوی دو      ۳. کمتر از  $n - 1$       ۴. کمتر از  $n$

۱۳- یک شرکت مولد وسایل برقی تاثیر سطح زیربنا در منزل های یک خانواده و نوع تهویه مطبوع مورد استفاده در منزل را در کل برق مصرفی طی ماههای با هوای گرم مورد رسیدگی قرار می دهد. فرض  $y$  کل برق مصرفی طی ماههای ژوئن، تا سپتامبر  $x_1$ ، مساحت منزل و  $x_p$  چهار نوع سیستم تهویه ۱- برون تهویه ۲- یونیت های پنجره دار ۳- پمپ گرما ۴- تهویه مرکزی باشد. چهار سطح عامل با سه متغیر نشانگر  $x_p, x_m, x_c$  بصورت زیر مدل سازی می شود.

نوع تهویه	$x_p$	$x_m$	$x_c$
بدون تهویه هوا	۰	۰	۰
پرینت های پنجره دار	۱	۰	۰
پمپ گرما	۰	۱	۰
تهویه صدای مرکزی	۰	۰	۱

اگر مدل رگرسیون بصورت  $y = \beta_0 + \beta_1 x_1 + \beta_p x_p + \beta_m x_m + \beta_c x_c$  باشد در حالتی که منزل تهویه نداشته باشد مدل کدام گزینه است؟

۱.  $y = \beta_0 + \beta_1 x_1 + \varepsilon$       ۲.  $y = (\beta_0 + \beta_1) + \beta_1 x_1 + \varepsilon$

۳.  $y = (\beta_0 + \beta_p) + \beta_1 x_1 + \varepsilon$       ۴.  $y = (\beta_0 + \beta_c) + \beta_1 x_1 + \varepsilon$

سری سوال: ۱ یک

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

تعداد سوالات: تستی: ۲۰ تشریحی: ۵

عنوان درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها، ریاضیات و کاربردها ۱۱۷۱۸۰

۱۴- فرض کنید در مدل رگرسیونی خطی ساده  $y = \beta_0 + x \beta_1 + \varepsilon$  با  $E(\varepsilon) = 0, \sigma^2 = Var(\varepsilon)$  غیر همبسته باشد در اینصورت:

$$Cov(\bar{y}, \hat{\beta}_1) = 0 \quad .2 \quad Cov(\hat{\beta}_0, \hat{\beta}_1) = 0 \quad .1$$

$$Cov(\hat{\beta}_0, \bar{y}) = 0 \quad .4 \quad Cov(\bar{y} + \hat{\beta}_0, \hat{\beta}_1) = 0 \quad .3$$

۱۵- در مدل رگرسیون چند متغیره  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$  نشان دهنده چیست؟

۱. عمر موثر وسیله

۲. سرعت برش

۳. عمق ارتفاع از مبدا صفحه رگرسیونی

۴. عمق برش

۱۶- کدامیک از گزینه های زیر همگی برای انتخاب بهترین متغیرهای رگرسیونی در یک مدل نمی باشد؟

۱. میانگین مربعات باقی مانده ها

۲. همه رگرسیون های ممکن

۳. حذف پیشرو

۴. روش قدم به قدم

۱۷- اگر در مدل  $y = X\beta + \varepsilon$  متغیرهای رگرسیونی را مرکزی شده و به طول واحد مقیاس سازی کنیم آن گاه  $X^1 X$  نشانگر کدام گزینه است؟

۱. ماتریس واریانس کواریانس

۲. ماتریس همبستگی

۳. ماتریس یکه I

۴. ماتریسی که تمام عضوهای آن یک است

۱۸- در یک مدل اگر بین متغیرهای رگرسیونی هم خطی شدید وجود داشته باشد کدام گزاره درست نمی باشد؟

۱. واریانس برآوردگرهای ضرایب رگرسیونی بزرگ می شود.

۲. کواریانس بین برآوردکننده های ضرایب رگرسیونی کوچک می شوند.

۳. نمونه های مختلف از سطوح یکسان X منجر به برآوردهای متفاوت می شود.

۴. همبستگی بین متغیرهای رگرسیونی و متغیر Y تغییری نمی کنند.

۱۹- کدامیک از گزینه های زیر روشهای بررسی هم خطی نمی باشد؟

۱. ماتریس همبستگی

۲. عوامل تورم واریانس

۳. مقادیر ویژه ماتریس  $X^1 X$

۴. جمع آوری داده های اضافی

۲۰- مدل ماتریسی مناسب رگرسیون چند متغیره کدام گزینه است؟

$$y = X\beta + \varepsilon \quad .1 \quad X = \beta + \varepsilon Y \quad .2 \quad X = \beta Y \quad .3 \quad y = X\beta \quad .4$$



تعداد سوالات: تستی: ۲۰ تشریحی: ۵

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

سری سوال: ۱ یک

عنوان درس: رگرسیون ۲

رشته تحصیلی/کد درس: آمار و کاربردها، ریاضیات و کاربردها ۱۱۷۱۸۰

## سوالات تشریحی

۱- در مدل رگرسیون چندمتغیره برآورد حداقل مربعات  $\beta$  را بدست آورید و نشان دهید که این برآوردگر نااریب برای  $\beta$  است؟

۲- نشان دهید از مدل رگرسیون چندگانه  $SS_R(\beta) = y'Hy$

۳- در مدل رگرسیون چندمتغیره  $y = x\beta + \varepsilon$  با فرض آنکه  $\varepsilon$  دارای توزیع نرمال با میانگین صفر و واریانس  $\sigma^2$  باشد برای نقطه مفروض  $X_{\cdot 1}, \dots, X_{\cdot k}$  برای میانگین پاسخ یک فاصله اطمینان بدست آورید.

۴- توابع  $y = \beta_0 + \beta_1 \log x$ ,  $y = \beta_0 + \beta_1 \left(\frac{1}{x}\right) + \varepsilon$  و  $y = \frac{x}{\beta_0 x - \beta_1}$  را در نظر بگیرید اگر این توابع قابلیت خطی شدن را دارند تبدیل لازم و شکل خطی را برای هر کدام بدست آورید.

۵- نشان دهید در مدل  $y = x\beta + \varepsilon$ ,  $v(\hat{y}) = \sigma^2 H$