

یاد خدا آرا بخش ولایت

موسسه
علوم
پزشکی

وزارت بهداشت، درمان و آموزش پزشکی
معاونت آموزشی
دبیرخانه شورای آموزش علوم پایه پزشکی، بهداشت و تخصصی
مرکز سنجش آموزش پزشکی

سوالات آزمون ورودی دکتری تخصصی (Ph.D)
سال تحصیلی ۹۴-۹۳

رشته: آمار زیستی

تعداد سوالات:	۱۰۰
زمان:	۱۵۰ دقیقه
تعداد صفحات:	۱۵

مشخصات داوطلب

نام:

نام خانوادگی:

داوطلب عزیز

لطفاً قبل از شروع پاسخگویی،

دفترچه سوالات را از نظر تعداد صفحات به دقت مورد بررسی قرار داده
و در صورت وجود هرگونه اشکال به مسئولین جلسه اطلاع دهید.

توجه: استفاده از ماشین حساب مجاز نمی باشد.

آمار زیستی



۱- اگر $X_1, X_2, \dots, X_n$ نمونه‌ای تصادفی و مستقل از توزیع نرمال استاندارد باشد و $T = X_1^2 + X_2^2 + X_3^2 + X_4^2$ کدام یک از متغیرهای زیر توزیع t دارد؟

(د) $\frac{5X_1}{\sqrt{T}}$

(ج) $\frac{2X_1}{\sqrt{T}}$

(ب) $\frac{X_1}{\sqrt{T}}$

(الف) $\frac{3X_1}{\sqrt{T}}$

۲- برآوردکننده گشتاوری پارامتر θ در چگالی زیر عبارتست از:

$f(x) = \begin{cases} (\theta+1)x^\theta, & 0 < x < 1 \\ 0 & \text{سایر جاها} \end{cases}$

(د) $\frac{1-\bar{X}}{1-2\bar{X}}$

(ج) $\frac{1-2\bar{X}}{1-\bar{X}}$

(ب) $\frac{2\bar{X}-1}{1-\bar{X}}$

(الف) $\frac{1-\bar{X}}{2\bar{X}-1}$

۳- اگر X دارای چگالی احتمال $f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$ برای $\beta > 0$ و $X > 0$ باشد، چگالی متغیر تصادفی $Y = x^\alpha$ برای $\alpha > 0$ و $y > 0$ عبارتست از:

(د) $\frac{\beta}{\alpha} y^{\alpha-1} e^{-\frac{y^\alpha}{\alpha}}$

(ج) $\frac{\alpha}{\beta} y^{\beta-1} e^{-\frac{y^\alpha}{\beta}}$

(ب) $\frac{\beta}{\alpha} y^{\alpha-1} e^{-\frac{y^\alpha}{\beta}}$

(الف) $\frac{\alpha}{\beta} y^{\alpha-1} e^{-\frac{y^\alpha}{\beta}}$

۴- تابع چگالی توأم $f(x,y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{سایر جاها} \end{cases}$ را در نظر بگیرید. احتمال $P[x+y < 1]$ عبارتست از:

(د) $\frac{1}{4}$

(ج) $\frac{1}{2}$

(ب) $\frac{1}{3}$

(الف) $\frac{2}{3}$

۵- اگر توزیع صفتی نامشخص و با واریانس متناهی باشد، فاصله $\mu \pm 2\sigma$ حداقل شامل چند درصد داده‌ها است؟

(د) ۹۷/۵٪

(ج) ۹۰٪

(ب) ۷۵٪

(الف) ۹۵٪

۶- اگر $X_1, X_2, \dots, X_n$ یک نمونه تصادفی مستقل از توزیع نرمال $N(\theta, 1)$ باشد، کران پائین کرامر-رائو برای واریانس برآوردگرهای نااریب θ^2 عبارتست از:

(د) $\frac{\theta^2}{n}$

(ج) $\frac{4\theta^2}{n}$

(ب) $\frac{4\theta}{n}$

(الف) $\frac{2\theta^2}{n}$

۷- برای آزمون $H_0: \sigma^2 \geq 4$ در مقابل $H_1: \sigma^2 < 4$ با مجهول بودن میانگین، کدام آماره آزمون مناسب است؟

(د) $\frac{(n-1)S^2}{4}$

(ج) $\frac{S^2}{4(n-1)}$

(ب) $\frac{nS^2}{4}$

(الف) $\frac{S^2}{4n}$

۸- X و Y دو متغیر تصادفی پیوسته و مستقل هستند که میانه هر دو صفر است، در این صورت $P(XY > 0)$ برابر است با:

(د) $\frac{2}{3}$

(ج) $\frac{1}{3}$

(ب) $\frac{1}{2}$

(الف) $\frac{1}{4}$



۹- اگر چگالی توأم X و Y به صورت $f(x,y) = \begin{cases} 12xy(1-y) & , 0 < x < 1, 0 < y < 1 \\ 0 & \text{سایر جاها} \end{cases}$ باشد، چگالی احتمال متغیر تصادفی $Z = XY^2$ عبارتست از:

الف) $g(z) = \begin{cases} 6+6z-12\sqrt{z} & , 0 < z < 1 \\ 0 & \text{سایر جاها} \end{cases}$ (الف)

ب) $g(z) = \begin{cases} 6z-12\sqrt{z} & , 0 < z < 1 \\ 0 & \text{سایر جاها} \end{cases}$ (ب)

ج) $g(z) = \begin{cases} 6-6z+12\sqrt{z} & , 0 < z < 1 \\ 0 & \text{سایر جاها} \end{cases}$ (ج)

د) $g(z) = \begin{cases} 6+6z+12\sqrt{z} & , 0 < z < 1 \\ 0 & \text{سایر جاها} \end{cases}$ (د)

۱۰- اگر X_1, X_2, X_3 ، یک نمونه تصادفی مستقل از هم، از توزیع برنولی با پارامتر θ باشند. کدام یک از آماره های ذیل برای θ آماره بسنده نیست؟

الف) $X_1 + 2X_2 + X_3$ (الف) ب) $X_1 + 5X_2 + X_3$ (ب) ج) $X_1 + 3X_2 + X_3$ (ج) د) $X_1 + 4X_2 + X_3$ (د)

۱۱- X متغیر تصادفی پیوسته دارای تابع توزیع $F(x)$ است، میانگین متغیر تصادفی $Y = (F(x))^2$ عبارتست از؟

الف) $\frac{2}{3}$ (الف) ب) $\frac{1}{3}$ (ب) ج) $\frac{1}{4}$ (ج) د) $\frac{1}{12}$ (د)

۱۲- اگر $X_1, X_2, \dots, X_n$ یک نمونه تصادفی n تایی از چگالی $f(x) = \theta x^{\theta-1}$; $0 < x < 1$, $\theta > 0$ باشد، برآورد گشتاوری θ عبارتست از:

الف) $\frac{1+\bar{X}}{\bar{X}}$ (الف) ب) $\frac{1-\bar{X}}{\bar{X}}$ (ب) ج) $\frac{\bar{X}}{1+\bar{X}}$ (ج) د) $\frac{\bar{X}}{1-\bar{X}}$ (د)

۱۳- اگر $X_1, X_2, \dots, X_n$ یک نمونه تصادفی مستقل از هم از یک توزیع برنولی با پارامتر P باشد، یک برآوردگر نااریب واریانس این توزیع عبارتست از:

الف) $X_1^2 - X_2^2 X_3^2$ (الف) ب) $X_1 + X_2 X_3$ (ب) ج) $X_1 - X_2 X_3$ (ج) د) $X_1^2 + X_2^2 X_3^2$ (د)

۱۴- در یک توزیع نرمال با میانگین معلوم μ و واریانس مجهول σ^2 ، برآورد حداکثر درستنمایی σ^2 عبارتست از:

الف) $\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$ (الف) ب) $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$ (ب)

ج) $\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$ (ج) د) $\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2$ (د)

۱۵- اگر X دارای توزیع دو جمله‌ای با پارامترهای n و p باشد، برآوردگر مینیمکس p چقدر است؟

الف) $\frac{1}{4}$ (الف) ب) $\frac{1}{3}$ (ب) ج) $\frac{1}{2}$ (ج) د) $\frac{1}{5}$ (د)

۱۶- اگر $X_1, X_2, \dots, X_n$ نمونه iid از چگالی $f(x) = \frac{\theta}{x^{\theta+1}}$; $\theta > 0$; $x \geq 1$ باشد، یک آماره بسنده θ عبارتست از:

الف) $\sum_{i=1}^n X_i$ (الف) ب) $\prod_{i=1}^n X_i$ (ب) ج) $\prod_{i=1}^n X_i^2$ (ج) د) $\sum_{i=1}^n X_i^2$ (د)

۱۷- برای یک توزیع دوجمله‌ای با پارامترهای $n=3$ و p ، برای یک آزمون فرض صفر $p = \frac{1}{2}$ در مقابل $H_1: p = \frac{2}{3}$ ، با ناحیه $R = \{x: x < 1\}$ مقدار خطای نوع دوم عبارتست از:

(د) $\frac{26}{27}$

(ج) $\frac{8}{27}$

(ب) $\frac{20}{27}$

(الف) $\frac{12}{27}$

۱۸- اگر X دارای توزیع نرمال با میانگین μ و واریانس ۱ باشد، برآورد UMVUE برای μ^2 عبارتست از:

(د) $\bar{X}^2 - \frac{1}{n}$

(ج) $\bar{X}^2 + \frac{1}{n}$

(ب) $\bar{X}^2$

(الف) $\bar{X} + \frac{1}{n}$

۱۹- اگر $X|Y$ دارای توزیع دوجمله‌ای با پارامترهای Y و P بوده و Y دارای توزیع پواسن با پارامتر λ باشد، EX عبارتست از:

(د) $\frac{\lambda}{P}$

(ج) λP

(ب) P

(الف) λ

۲۰- نمونه تصادفی X_1, X_2, X_3 از متغیر تصادفی X با میانگین μ و واریانس متناهی می‌باشد، کارآیی نسبی برآوردکننده $\frac{1}{4}(X_1 + 2X_2 + X_3)$ نسبت به $\bar{X}$ عبارتست از:

(د) $\frac{4}{3}$

(ج) $\frac{3}{4}$

(ب) $\frac{8}{9}$

(الف) $\frac{9}{8}$

روش‌های آمار زیستی

۲۱- در یک تحقیق سه کلاس انتخاب شده و در هر کلاس به طور تصادفی یکی از سه روش آموزشی تدریس شده است. اگر مدل رگرسیونی بین نمرات بعد از آموزش (y_i) با قبل از آموزش (x_i) در هر گروه به صورت $y_i = \alpha_j + 2x_i$ باشد $(j=1,2,3)$ کدام روش برای آنالیز مناسب‌تر است؟

(الف) آنالیز واریانس یک‌طرفه $d_i = y_i - x_i$ بین سه گروه

(ب) آنالیز کوواریانس نمرات y_i با تعدیل نسبت به x_i

(ج) آنالیز واریانس یک‌طرفه y_i ها بین سه گروه

(د) آنالیز کوواریانس x_i ها با تعدیل نسبت به y_i

۲۲- در رگرسیون لجستیک کدام گزینه درست می‌باشد؟

(الف) ضریب هر متغیر مستقل، تغییر یک واحد odds را برای آن متغیر نشان می‌دهد.

(ب) توزیع خطای مدل نرمال است.

(ج) واریانس خطا برای پاسخ‌های مختلف ثابت است.

(د) ضریب هر متغیر مستقل، لگاریتم تغییر یک واحد odds را برای آن متغیر نشان می‌دهد.

۲۳- مدل $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ که در آن $y_i = 0, 1$ است را در نظر بگیرید. اگر x ثابت باشد، حداکثر چند مقدار مختلف می‌توان برای ε_i در نظر گرفت؟

(الف) ۱

(ب) ۲

(ج) به تعداد مشاهدات

(د) به تعداد نصف مشاهدات

کدامیک از ملاک های زیر تغییرات حاصل از حذف یک فرد را بر برازندگی کلی مدل مناسب تر ارزیابی می کند؟

الف) Likelihood Ratio

ب) DF Beta

ج) DF Fit

د) Leverage

۲۵- در یک مدل رگرسیون خطی بین Y (متغیر پاسخ) و X_1 و X_2 (متغیرهای مستقل) برآورد ضرایب β_1 و β_2 از نظر آماری معنی دار شده اند اما ضریب تعیین مدل به اندازه کافی بزرگ نیست، بیان کدام گزینه زیر مناسب تر است؟

الف) این مدل فقط برای بیان تاثیر متغیرهای مستقل بر پاسخ مناسب است.

ب) این مدل برای پیش بینی متغیر پاسخ مناسب است.

ج) این مدل هم برای بیان تاثیر متغیرهای مستقل بر پاسخ و هم برای پیش بینی آن مناسب است.

د) چنین مدلی ارزش آماری ندارد.

۲۶- یک طرح آشیانه ای را با دو عامل ثابت در نظر بگیرید که عامل B درون A (nested) می باشد. در صورتی که تعداد سطوح A و B برابر و مساوی ۳ و تعداد تکرار ۵ باشد، درجه آزادی خطا برابر است با:

الف) ۸ (ب) ۱۲ (ج) ۱۶ (د) ۳۶

۲۷- در مدل رگرسیون خطی چندگانه $E(Y|X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$ با افزایش یک واحد در متغیر مستقل X_1 ، به طور متوسط Y چقدر تغییر خواهد کرد؟

الف) $\beta_0 + \beta_1 + \beta_2 X_2$ (ب) β_1 (ج) $\beta_0 + \beta_1$ (د) $\beta_1 + \beta_2 X_2$

۲۸- کدامیک از گزینه های زیر در مورد رگرسیون پواسنی صحیح است؟

الف) مدل مناسب برای آن خطی است.

ب) در این مدل معمولا متغیر پاسخ شمارشی برای تعداد آزمایش زیاد و وقایع نادر است.

ج) متغیر پاسخ آن دو مقداری است.

د) خطاها دارای توزیع نرمال هستند.

۲۹- یک طرح تک عاملی با اندازه های تکراری بر روی عامل ثابت A را در نظر بگیرید. اگر واریانس مقادیر پاسخ برابر ۶۴ و کوواریانس مشاهدات برابر ۴۹ باشد، مقدار همبستگی درونی (دو مشاهده برای هر فرد) چقدر است؟

الف) $\frac{49}{64}$ (ب) $\frac{15}{64}$ (ج) $\frac{7}{8}$ (د) $\frac{1}{8}$

۳۰- میانگین، میانه و انحراف معیار مدت بستری در بیمارستان برای یک گروه از بیماران به ترتیب ۶ و ۴ و ۸ روز و در گروه دوم به ترتیب ۸ و ۵ و ۱۰ روز بوده است. کدامیک از آزمون های زیر برای مقایسه مدت بستری دو گروه مناسب تر است؟

الف) ویلکاکسون برای رتبه های علامت دار

ب) من-ویتنی

ج) تی مستقل

د) آنالیز واریانس یک طرفه

۳۱- فرض کنید یکصد هزار داوطلب ورود به دانشگاه براساس نمرات آزمون به ترتیب نزولی مرتب شده‌اند و سپس در فهرست مرتب شده هر صد نفر متوالی به عنوان یک خوشه در نظر گرفته شده است. (هزار خوشه صد نفری) می‌خواهیم نمونه‌ای به حجم $n=1000$ برای برآورد میانگین بهره هوشی انتخاب کنیم کدامیک از روش‌های نمونه‌گیری زیر از دقت بالاتری برخوردار است؟

(الف) انتخاب بیست خوشه و سپس انتخاب پنجاه نفر از هر خوشه

(ب) نمونه‌گیری تصادفی ساده از فهرست مرتب شده

(ج) انتخاب ده خوشه و بررسی تمام افراد درون خوشه

(د) نمونه‌گیری سیستماتیک از فهرست مرتب شده

۳۲- یک طرح دو عاملی با اندازه‌های تکراری بر روی یک عامل، معادل کدام طرح است؟

(الف) بلوک‌های ناقص تصادفی

(ب) آشیانه‌ای (Nested)

(ج) کرت‌های خرد شده

(د) بلوک‌های کامل تصادفی

۳۳- ضریب همبستگی کندال بین x و $\ln(x)$ برای $x > 0$ برابر است با:

(الف) ۱ (ب) صفر (ج) ۰/۵ (د) -۱

۳۴- در مدل رگرسیونی $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ عامل تورم واریانس (VIF) بزرگتر از ۱۰ شده است. کدامیک از

گزینه‌های زیر در مورد ضریب همبستگی پیرسن (r) بین دو متغیر تصادفی X_1 و X_2 درست است؟

(الف) $r = 0.90$ (ب) $0.17 < r < 0.19$ (ج) $|r| > 0.190$ (د) $r = -0.190$

۳۵- می‌خواهیم اثر آموزش را بر کیفیت زندگی بیماران مبتلا به MS در دو گروه شاهد و مداخله پس از ۳ ماه ارزیابی

نماییم. چنانچه میانگین کیفیت زندگی در شروع مطالعه در دو گروه شاهد و مداخله یکسان نباشد، کدام روش

مناسب است؟

(الف) t مستقل براساس مقایسه نتایج ماه سوم

(ب) t زوجی

(ج) آنالیز کوواریانس

(د) بلوک کامل تصادفی

۳۶- جامعه‌ای از دو طبقه به حجم‌های $N_1=3$ و $N_2=6$ تشکیل شده است. می‌خواهیم $n=3$ نمونه به روش تصادفی ساده

انتخاب کنیم. احتمال اینکه نمونه انتخاب شده دقیقاً متناسب با سایز باشد، چقدر است؟

(الف) $\frac{3}{28}$ (ب) $\frac{15}{28}$ (ج) $\frac{5}{28}$ (د) $\frac{1}{3}$

۳۷ - در یک کار آزمایی بالینی با دو گروه مستقل، تعداد نمونه براساس مطالعات پیشین و $\alpha=0.05$ محاسبه شده است. پس از انجام تحقیق توان آزمون به کمتر از مقدار فرض شده قبلی رسیده است. کدام گزینه صحیح است؟ (σ^2 واریانس و d اندازه اثر است)

الف) کاهش برآورد $\sigma^2 d^2$ نسبت به مقدار مفروض

ب) افزایش برآورد $\frac{d^2}{\sigma^2}$ نسبت به مقدار مفروض

ج) افزایش برآورد $\sigma^2 d^2$ نسبت به مقدار مفروض

د) کاهش برآورد $\frac{d^2}{\sigma^2}$ نسبت به مقدار مفروض

۳۸ - حد بالای فاصله اطمینان ۹۰٪ واریانس خطا در یک طرح آنالیز واریانس یک طرفه با اثرات تصادفی چقدر است در صورتی که درجه آزادی خطا برابر ۱۵ و $MSE=75$ ، $\chi^2(0.05, 15) \approx 24$ و $\chi^2(0.95, 15) \approx 25$ باشد؟

الف) ۳ (ب) ۴۵ (ج) ۱۰۵ (د) ۱۵۰

۳۹ - در یک طرح آشیانه‌ای B درون A (nested) می‌باشد. تعداد سطوح A برابر ۴ و تعداد سطوح B برابر ۳ و تعداد تکرار ۵ می‌باشد. اگر $SS(A)=15$ ، $SS(B)=20$ و $SS(AB)=12$ باشد، مقدار $MSB(A)$ چقدر است؟

الف) ۴ (ب) ۷ (ج) ۱۶ (د) ۳۲

۴۰ - طرح با اندازه‌گیری مکرر برای بررسی اثر درمانی A که در دو سطح می‌باشد را روی ۱۰ بیمار انجام داده‌ایم. جهت آزمون اختلاف تاثیر دو سطح درمانی در سطح $\alpha=0.05$ نقطه بحرانی کدام است؟

الف) $F_{0.95}$ با ۸ درجه آزادی

ب) $F_{0.975}$ با ۹ درجه آزادی

ج) $t_{0.95}$ با ۹ درجه آزادی

د) $t_{0.975}$ با ۹ درجه آزادی

۴۱ - اگر متغیر تصادفی X دارای توزیع پواسن با میانگین λ_1 و متغیر تصادفی Y دارای توزیع پواسن با میانگین λ_2 باشد، برای کمیت تصادفی X-Y می‌توان گفت:

الف) دارای توزیع پواسن با میانگین $\lambda_1 - \lambda_2$ است.

ب) دارای واریانس $\lambda_1 + \lambda_2$ است.

ج) دارای واریانس $\lambda_1 - \lambda_2$ است.

د) دارای توزیع پواسن با واریانس $\lambda_1 + \lambda_2$ است.

۴۲ - در یک رگرسیون خطی $\hat{y} = 2 + 3x$ اگر $SSX=10$ و $SST=360$ آنگاه ضریب تعیین برابر است با:

الف) $\frac{1}{2}$ (ب) $\frac{1}{4}$ (ج) $\frac{2}{3}$ (د) $\frac{1}{36}$

۴۳ - یک رگرسیون خطی شامل سه متغیر توضیحی x_1 و x_2 و x_3 می‌باشد. کدامیک از گزینه‌ها صحیح نیست؟

الف) $SSR(x_1, x_2, x_3) = SST(x_1, x_2, x_3) - SSE(x_1, x_2, x_3)$

ب) $SSR(x_1, x_2, x_3) = SSR(x_1) + SSR(x_2, x_3 | x_1)$

ج) $SSR(x_1, x_2, x_3) = SSR(x_1) + SSR(x_2 | x_1) + SSR(x_3 | (x_1, x_2))$

د) $SSR(x_1, x_2, x_3) = SSE(x_1) + SSE(x_2 | x_1) + SSE(x_3 | x_1, x_2)$

۴۴- یک طرح آزمایشی با دو عامل ثابت و با اندازه‌های تکراری بر روی هر دو عامل به صورت کامل اجرا می‌نماییم. اگر عامل A در دو سطح و عامل B در سه سطح و تعداد افراد مورد آزمایش ۱۰ نفر باشد، درجه آزادی خطا برابر است با:

الف) ۳ (ب) ۹ (ج) ۶ (د) ۱۸

۴۵- در یک رگرسیون چندگانه با دو متغیر مستقل X_1 و X_2 مقادیر b_1, b_2 برآوردهای β_1 و β_2 در مدل $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i$ می‌باشد. اگر بین X_1 و X_2 همخطی شدید وجود داشته باشد، آنگاه در مورد مجموع مربعات می‌توان گفت:

الف) $SSE(b_1, b_2) > SSE(b_1)$

ب) $SSR(b_1) < SSR(b_1 | b_2)$

ج) $SSR(b_1) > SSR(b_1 | b_2)$

د) $SSR(b_1, b_2) < SSR(b_1)$

۴۶- در یک رگرسیون خطی چندگانه با ۵ متغیر توضیحی و حجم نمونه ۳۱ اگر ضریب تعیین $R^2 = 0.80$ باشد، ضریب تعیین تعدیل شده چقدر است؟

الف) $\frac{10}{13}$ (ب) $\frac{19}{25}$ (ج) $\frac{11}{13}$ (د) $\frac{20}{25}$

۴۷- در مدل $P(x) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2}}$ وقتی X_1 پیوسته و $X_2 = 0, 1$ باشد، آنگاه نسبت شانس X_2 برابر است با:

الف) $\exp(\beta_2)$ (ب) $\exp(\beta_1)$ (ج) $\exp(\beta_1 \beta_2)$ (د) $\exp(\beta_1 + \beta_2 X_1)$

۴۸- در مدل رگرسیون خطی $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ که Y دارای توزیع نرمال است و $\hat{\beta}_1$ برآورد β_1 باشد آنگاه:

الف) $P(\hat{\beta}_1 > \beta_1) < \frac{1}{2}$ (ب) $P(\hat{\beta}_1 > \beta_1) = 0$ (ج) $P(\hat{\beta}_1 > \beta_1) = \frac{1}{2}$ (د) $P(\hat{\beta}_1 > \beta_1) > \frac{1}{2}$

۴۹- یک توزیع کای - دو با درجه آزادی n در چه حالتی به توزیع نمایی ساده تبدیل می‌شود؟

الف) $n=1$ (ب) $n=2$ (ج) $n=4$ (د) $n=3$

۵۰- در مدل‌های رگرسیونی خطی ساده که $\hat{Y}$ برآورد رگرسیونی Y است، ضریب همبستگی بین X و Y را با r و $\hat{Y}$ را با r' نمایش می‌دهیم. کدام یک از گزینه‌های زیر صحیح است؟

الف) $r > r'$ (ب) $r < r'$ (ج) $r = r'$ (د) $r' = r'$

۵۱- یک طرح دو عاملی کامل را در نظر بگیرید که A ثابت و B تصادفی می‌باشد. تعداد سطوح A برابر ۲ و تعداد سطوح B برابر ۳ و تعداد تکرار ۴ می‌باشد، در صورتی که $SSB=20$ و $MSE=2$ باشد، برآورد ناریب واریانس عامل B چقدر است؟

الف) ۱ (ب) ۱/۵ (ج) ۲ (د) ۳

۵۲- مدل رگرسیون چندگانه $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ را به داده های حاصل از ۱۸ نفر برازش نموده ایم. با توجه به مدل کاهش یافته $y = \beta_0 + \beta_1 X_1$ اگر $SSE(مدل کامل) = ۱۵$ و $SSE(مدل کاهش یافته) = ۲۰$ باشد، مقدار F آزمون $H_0: \beta_2 = 0$ چقدر است؟

(الف) ۲

(ب) ۳

(ج) ۴

(د) ۵

۵۳- در یک نمونه تصادفی از توزیع توام X و Y ضریب خط رگرسیونی $Y|X$ را با b و ضریب $X|Y$ را با b' نشان میدهیم. در اینصورت اگر ضریب همبستگی پیرسون بین X و Y را با r نمایش دهیم. کدام گزینه صحیح است؟

(الف) $r^2 = bb'$ (ب) $r^2 = \sqrt{bb'}$ (ج) $r^2 = (bb')^2$ (د) $r^2 = b + b'$

۵۴- در مدل رگرسیون $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ اگر ضریب همبستگی پیرسون بین X_1 و X_2 برابر $\frac{1}{3}$ باشد، آنگاه کدام گزینه زیر عامل تورم واریانس را نشان می دهد؟

(الف) $1/66$ (ب) $1/33$

(ج) ۲

(د) ۴

۵۵- در مدل رگرسیون خطی چندگانه $E(Y|X) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$ آماره F مدل معنی دار شده است. اما آزمون های مربوط به ضرایب رگرسیونی هیچکدام معنی دار نشده اند. دلیل این امر کدامیک از موارد زیر است؟

- (الف) حجم نمونه کم است
(ب) ممکن است خود همبستگی بین مشاهدات وجود داشته باشد.
(ج) ناهمگنی واریانس وجود دارد.
(د) همخطی بین متغیرهای مستقل وجود دارد.

تحلیل چند متغیره

۵۶- اگر $X_1, X_2, \dots, X_n$ نمونه های تصادفی و مستقل از توزیع با میانگین μ و ماتریس کوواریانس Σ باشد و $S = \frac{1}{n} \sum_{i=1}^n X_i X_i' - \bar{X} \bar{X}'$ آنگاه $E(S)$ عبارت است از:

(الف) $\frac{1}{n} \Sigma$ (ب) $\frac{n}{n-1} \Sigma$ (ج) $\frac{1}{n-1} \Sigma$ (د) $\frac{n-1}{n} \Sigma$

۵۷- در مولفه های اصلی (Principal Components)

- (الف) ماتریس واریانس کوواریانس (S) می تواند منفرد باشد.
(ب) ضرایب مولفه های اصلی متاثر از بزرگ یا کوچک بودن واریانس متغیرهای اولیه نیستند.
(ج) مولفه های اصلی مبتنی بر ماتریس همبستگی (R) یکتا هستند.
(د) مولفه های اصلی نسبت به تغییر مقیاس متغیرها پایا هستند.

۵۸- کدامیک از موارد زیر جزء رهیافت های ارزیابی نرمالیتی چند متغیره نیست؟

- (الف) آزمون نرمال بودن ترکیب های غیر خطی متغیرهای اولیه
(ب) آزمون آماری چند متغیره کای دو
(ج) آزمون نرمال بودن ترکیب های خطی متغیرهای اولیه
(د) ارزیابی نرمالیتی تک تک متغیرها

۵۹- در آزمون بررسی تساوی دو بردار وابسته، کدام گزینه مناسب است؟

- الف) استفاده از آزمون T^2 چند متغیره هتلینگ دو نمونه مستقل به جای T^2 هتلینگ زوجی، موجب کاهش توان آماری میشود.
- ب) در آزمون T^2 هتلینگ زوجی چند متغیره، نرمال بودن تک متغیره تفاضل ها کفایت میکند.
- ج) در آزمون T^2 هتلینگ زوجی چند متغیره، برقراری شرط نرمال بودن چند متغیره برای مقادیر قبل از مداخله کفایت می کند.
- د) استفاده از آزمون T^2 چند متغیره هتلینگ دو نمونه مستقل به جای T^2 هتلینگ زوجی، موجب افزایش توان آماری میشود.

۶۰- در تحلیل عاملی (Factor Analysis)

- الف) روش حداکثر درست نمایی (ML) جهت استخراج عوامل، تنها روش توصیه شده است.
- ب) واریانس های مشترک یکتا نیستند.
- ج) واریانس های مشترک یکتا هستند.
- د) عوامل استخراجی در همه روش های استخراج از نظر تبیین تغییرپذیری داده ها همواره از سیر نزولی برخوردار است.

۶۱- در تحلیل همبستگی کانونی کدامیک از گزینه های زیر صحیح است؟

- الف) متغیرهای کانونی دارای واریانس یک هستند.
- ب) اگر فرض صفر $\sum_{xy} = 0$ رد شود آنگاه همه ضرایب همبستگی کانونی مخالف صفر خواهند بود.
- ج) ضرایب استاندارد شده در متغیرهای کانونی (Canonical Variables) در تفسیر نقش متغیرها ماهیت چندمتغیره ندارند.
- د) اطلاعات موجود در همبستگی های کانونی مراتب مختلف از هم مستقل و مجزا هستند.

۶۲- اگر $X_1, X_2, \dots, X_p$ نمونه ای مستقل از یک جامعه نرمال چند متغیره $N_p(\mu, \Sigma)$ باشد، آنگاه:

- الف) $\bar{X} \sim N(\mu, \frac{\Sigma}{n})$
- ب) $S \sim W_p(19, \Sigma)$ (ماتریس واریانس - کواریانس و W نشانگر توزیع ویشارت است)
- ج) $\bar{X} \sim N(\mu, \Sigma)$
- د) $S \sim W(20, \Sigma)$ (ماتریس واریانس - کواریانس و W نشانگر توزیع ویشارت است)

۶۳- تابع رده بندی (Classification Function) فیشر

- الف) در حالت $\sum_1 \neq \sum_2$ و به فرض نرمال بودن داده ها بهینه عمل می کند.
- ب) با تابع ممیزی Discriminant Function یکسان است.
- ج) در حالت $P_1 = P_2$ (احتمال های پیشین) و نرمال بودن داده ها بهینه عمل می کند.
- د) می تواند جوامع با هر نوع روند در توزیع داده ها را به بهینه ترین حالت تمایز دهد.

۶۴- در ارتباط با شاخص های تبیین تغییرپذیری چند متغیره کدام گزینه صحیح است؟

- الف) واریانس تعمیم یافته هم تغییری متغیرها را لحاظ نمی کند.
- ب) واریانس تعمیم یافته (Generalized variance) نمی تواند صفر شود.
- ج) واریانس تعمیم یافته (Generalized variance) در مقایسه با مجموع واریانس (Trace S) پدیده همخطی را دقیق تر منعکس می کند.
- د) تغییرپذیری موجود در داده های چند متغیره در قالب یک عدد قابل خلاصه کردن نیست.

... (Discriminant) میزانی

الف) جهت توابع ممیزی یکسان است.

ب) متغیرهای مورد بررسی نمیتوانند هر مقیاسی داشته باشند.

ج) همیشه نمی توان زیر مجموعه ای از متغیرهای اولیه را پیدا کرد که قدرت ممیزی آن با متغیرهای اولیه یکسان باشد.

د) هم خطی یا مشاهدات پرت تاثیر جزئی بر عملکرد توابع ممیزی دارد.

۶۶- اگر بردارهای μ_1 و μ_2 میانگین دو جامعه مورد بررسی باشند، بعد از رد فرضیه صفر $H_0: \mu_1 = \mu_2$ توسط آزمون T^2 هتلینگ مهمترین آزمون پسین کدام است؟

الف) آزمون F یا t جزئی

ب) آزمون های t تک متغیره

ج) آزمون ترکیب های دو تایی

د) آزمون توام t تک متغیره و دو تایی

۶۷- اگر S (ماتریس واریانس-کواریانس) و R (ماتریس همبستگی) و D ماتریس قطری انحراف معیارها باشند، آنگاه:

الف) $E(S) \neq \sum$ (ماتریس واریانس - کواریانس جامعه)ب) $E(R) \neq P$ (ماتریس همبستگی جامعه است)ج) $R = D^{-1}SD^{-1}$ د) $S = D^{1/2}RD^{1/2}$

۶۸- در مورد فاصله چند متغیره ماهالانوبیس کدامیک از گزینه های زیر صادق نیست؟

الف) همبستگی متغیرها را حذف می کند.

ب) دو متغیری که همبستگی بالاتری نسبت به دو متغیری که همبستگی کمتر دارند نقش بیشتری ایفا می کنند.

ج) متغیرها را نسبت به یک واریانس یکسان استاندارد می کند.

د) متغیری که واریانس بزرگتر دارد، وزن کمتری می گیرد.

۶۹- در تحلیل خوشه بندی کدامیک از گزینه های زیر در مورد روش های پیوند (Linkage) درست است؟

الف) در روش پیوند Centroid فاصله بین دو خوشه A و B به صورت فاصله اقلیدسی بین بردارهای میانگین دو خوشه تعریف می شود.

ب) در روش Average Linkage فاصله بین دو خوشه A و B که به ترتیب دارای n_A و n_B عضو هستند عبارت است از میانگین مجموع $n_A + n_B$ است.

ج) در روش Ward دو خوشه ای به هم می پیوندند که منجر به بیشترین افزایش SSE بین خوشه ها شوند.

د) در روش Ward دو خوشه ای به هم می پیوندند که منجر به بیشترین کاهش در SSE درون خوشه ها شوند.

۷۰- همبستگی کانونی (Cononical Correlation) دو مجموعه متغیر X و Y

الف) می تواند عددی منفی باشد.

ب) رابطه دو مجموعه را اندازه می گیرد.

ج) رابطه خطی دو مجموعه را اندازه می گیرد.

د) رابطه خطی ترکیب خطی دو مجموعه را اندازه می گیرد.

۷۱- اگر $S_1(t)$ تابع بقاء برای سیگاری ها و $S_2(t)$ تابع بقاء برای غیر سیگاری ها باشد و $S_2(t) = S_1(0.75t)$ باشد، در این صورت:

- الف) بقاء ۸۰ ساله غیر سیگاری ها معادل ۴۰ سال سیگاری ها است.
- ب) فاکتور شتابنده در این مدل برای غیر سیگاری ها ۰/۷۵ سیگاری ها است.
- ج) فاکتور شتابنده در این مدل برای سیگاری ها ۱/۲۵ غیر سیگاری ها است.
- د) بقاء ۸۰ ساله غیر سیگاری ها معادل ۶۰ سال سیگاری ها است.

۷۲- کدامیک از گزاره های زیر نادرست است؟

- الف) شکنندگی مشترک مشابه اثر تضادفی در مدل های رگرسیون است.
- ب) شکنندگی مشترک در مدل COX همبستگی بین گروه ها را در نظر می گیرد.
- ج) مدل های شکنندگی غیر مشترک افراد را وابسته می گیرد.
- د) مدل های شکنندگی غیر مشترک افراد را از یکدیگر مستقل می گیرد.

۷۳- کدامیک از مدل های زیر را می توان بر حسب T یا $\ln(T)$ بیان نکرد؟

- الف) گمپرتز (ب) وایبل (ج) لگ- لجستیک (د) لگ- نرمال

۷۴- در یک رویکرد آزمون نکویی برازش در مدل خطرات متناسب کاکس، اگر فرض خطرات متناسب برای یک متغیر توضیحی خاص برقرار باشد، بنابراین:

- الف) باقیمانده های کاکس - اسنل با زمان بقاء همبستگی خواهند داشت.
- ب) باقیمانده های شونفلد برای متغیر توضیحی با زمان بقاء همبستگی نخواهد داشت.
- ج) باقیمانده های مارتینگل برای متغیر توضیحی با زمان بقاء همبستگی خواهند داشت.
- د) باقیمانده های شونفلد برای متغیر توضیحی با زمان بقاء همبستگی خواهند داشت.

۷۵- کدامیک از مدل های زیر اثر متغیر توضیحی نسبت به زمان بقا ضرب پذیر نمی باشد؟

- الف) وایبل (ب) لگ نرمال (ج) لگ- لجستیک (د) گمپرتز

۷۶- جهت بررسی مناسبت توزیع لگ - لجستیک در داده های بقاء کدامیک از موارد زیر استفاده می شود؟

- الف) رسم نمودار لگاریتم بخت شکست در مقابل لگاریتم زمان بقاء
- ب) رسم نمودار، لگاریتم منفی بخت شکست در مقابل زمان بقاء
- ج) رسم نمودار $\left|1 - \hat{S}(t)\right| \phi^{-1}$ در مقابل لگاریتم زمان بقاء
- د) رسم نمودار $\log \left| -\log \hat{S}(t) \right|$ در مقابل لگاریتم زمان بقا

۷۷- زمانی که در مدل کاکس توابع بقاء برای دو گروه که در مقابل زمان رسم شده است، همدیگر را قطع نکنند

- الف) همواره فرضیه خطرات متناسب برقرار است.
- ب) فرضیه خطرات متناسب ممکن است برقرار نباشد.
- ج) همواره فرضیه خطرات متناسب برقرار نیست.
- د) همواره نمی توان از مدل کاکس طبقه بندی شده استفاده کرد.



۷۸- می‌دانیم که تابع مخاطره وایبل به صورت $h(t) = \lambda t^{p-1}$ می‌باشد. اگر $\lambda = \exp(\alpha_0 + \alpha_1 E)$ و $P = \exp(\beta_0 + \beta_1 E)$

باشد که E برابر یک اگر مداخله داشته باشیم و صفر اگر مداخله نداشته باشیم. در این صورت:

(الف) تابع همواره دارای خاصیت AFT است.

(ب) تابع همواره دارای خاصیت PH است.

(ج) اگر $\beta_1 = 0$ باشد این مدل همواره PH و AFT است.

(د) در صورتی که $\alpha_1 = 0$ باشد این مدل همواره PH و AFT است.

۷۹- کدامیک از روش‌های زیر برای آنالیز داده‌های مکرر (Recurrent) در حالتی که وقایع متفاوت هستند، توصیه می‌شود؟

(د) Stratified CP

(ج) Marginal

(ب) CP

(الف) GAP Time

۸۰- در مدل لگ- لجستیک

(الف) همواره odds بقاء در طول زمان ثابت است.

(ب) همواره نسبت odds بقاء در طول زمان ثابت است.

(ج) همواره odds شکست در طول زمان ثابت است.

(د) همواره خاصیت PH را دارا است.

۸۱- اگر تابع مخاطره به صورت $h(t) = 3\lambda t^2$ باشد و مقدار λ برحسب وجود مداخله $E=1$ و غیرمداخله $E=0$ به صورت

$\lambda = \exp(2 + 1/5E)$ باشد، در این صورت تابع پایه (Base Line) این مدل عبارت است از:

(د) $\exp(2)t^2$

(ج) $3\exp(1.5)t^2$

(ب) $3t^2$

(الف) $3\exp(2)t^2$

۸۲- کدامیک از روش‌های زیر از مدل SC (Stratified COX) برای آنالیز داده‌های مکرر (Recurrent) به کار برده می‌شود؟

(د) همه موارد

(ج) Stratified CP

(ب) Gap Time

(الف) Marginal

۸۳- مقدار P (P-value) به دست آمده از Robust Standard Error در مقایسه با P به دست آمده از

Nonrobust Standard Error

(الف) همواره بزرگ‌تر است.

(ب) بعضی مواقع بزرگ‌تر و بعضی مواقع کوچک‌تر است.

(ج) همواره کوچک‌تر است.

(د) در صورتی که در CP به کار برده شود بزرگ‌تر است.

۸۴- $E(t)$ مواجهه با آلودگی هوا در روزهای مختلف می‌باشد اگر مقدار آن برابر صفر در حالتی که مواجهه ندارد و در

حالتی که مواجهه دارد برابر ۱ باشد. در این صورت $h(t) = h_0(t) \exp(\delta E(t))$

(الف) همواره PH است.

(ب) چون $HR(t)$ ثابت است پس PH است.

(ج) $HR(t)$ وابسته به زمان است.

(د) همواره AFT است.

۸۵- فرضیه‌های لگ نرمال کدام ویژگی موجود نیست؟

- الف) پارامتری بودن
- ب) فرضیه شانس تناسب (Proportional odds)
- ج) زمان شکست شتابنده (AFT)
- د) فرضیه خطرات تناسب (Proportional Hazard)

کار آزمایی بالینی

۸۶- منظور از اثر متقابل تیمار - دوره در یک کار آزمایی بالینی متقاطع چیست؟

- الف) اثر دوره
- ب) انتقال اثر دارو از مرحله اول به مرحله دوم
- ج) اثر خطای اندازه‌گیری
- د) اثر انتخاب ناصحیح نمونه‌ها

۸۷- کدام یک از موارد زیر از مزایای یک کار آزمایی بالینی متقاطع است؟

- الف) تعداد نمونه کمتر در گروه‌ها نسبت به گروه‌های موازی
- ب) تعداد نمونه بیشتر در گروه‌ها نسبت به گروه‌های موازی
- ج) عدم کنترل ویژگی فردی نسبت به گروه‌های موازی
- د) وجود اثر دوره - تیمار نسبت به گروه‌های موازی

۸۸- کدام عبارت زیر در مورد روش کمینه‌سازی (Minimization) صحیح است؟

- الف) همان روش تصادفی بلوکی است که در نمونه‌های کوچک کاربرد دارد.
- ب) روشی کاملاً تصادفی است و در کار آزمایی بالینی با تعداد نمونه کم کاربرد دارد.
- ج) همان روش تصادفی طبقه‌بندی شده است که در نمونه‌های بزرگ اجرا می‌شود.
- د) روشی نیمه تصادفی است و در کار آزمایی بالینی با تعداد نمونه کم کاربرد بیشتری دارد.

۸۹- نمونه‌ها در کار آزمایی بالینی مرحله اول عبارتند از:

- الف) افراد بیماری که به تصادف انتخاب می‌شوند.
- ب) افراد بیمار و داوطلب
- ج) افرادی سالمی که به تصادف انتخاب می‌شوند.
- د) افراد سالم و داوطلب

۹۰- در چه شرایطی به کارگیری تصادفی سازی طبقه‌ای شده بیشتر توصیه می‌شود؟

- الف) تعداد طبقات زیاد و افراد داخل طبقات کم باشد.
- ب) تعداد طبقات کم و افراد داخل طبقات زیاد باشد.
- ج) تعداد طبقات کم و افراد داخل طبقات ناهمگون باشند.
- د) تعداد طبقات زیاد و افراد داخل طبقات کم و ناهمگون باشند.

۹۱- در یک کارآزمایی بالینی مهمترین ویژگی برای استفاده از دارونما چیست؟

- (الف) شناسایی مثبت واقعی
- (ب) شناسایی منفی کاذب
- (ج) شناسایی مثبت کاذب
- (د) شناسایی منفی واقعی

۹۲- کدام یک از موارد زیر از مشکلات مهم یک کارآزمایی بالینی متقاطع است؟

- (الف) کنترل ویژگی های فردی در نمونه های کوچک تر نسبت به گروه های موازی
- (ب) داشتن تعداد نمونه کوچک تر در گروه ها نسبت به گروه های موازی
- (ج) کنترل ویژگی های فردی در گروه ها
- (د) احتمال ریزش بیمار در فاصله دو دوره (Wash-out)

۹۳- چرا اثر دوره - تیمار (انتقالی) در کارآزمایی بالینی متقاطع ایجاد می شود؟

- (الف) عدم انتخاب گروه کنترل
- (ب) کوتاه بودن دوره شستشو (Wash-out)
- (ج) کم بودن تعداد نمونه در گروه های مورد مطالعه
- (د) انتخاب ناصحیح بیمار در مطالعه

۹۴- در یک کارآزمایی بالینی دو گروهی با $\alpha = 0.05$ اگر بخواهیم سه مقایسه در هفته های اول و دوم و سوم انجام دهیم در این حالت حداکثر مقدار $p = \text{Value}$ برای معنی داری چقدر است؟

- (الف) ۰/۰۱۷ (ب) ۰/۰۲۵ (ج) ۰/۰۵۰ (د) ۰/۰۱۲۵

۹۵- در کارآزمایی بالینی به کارگیری تصادفی سازی طبقه بندی شده در کدام یک از موارد زیر توصیه می شود؟

- (الف) وقتی بخواهیم تعادل در اندازه نمونه در گروه ها تحت مطالعه برقرار کنیم.
- (ب) ویژگی های تاثیرگذار بر پاسخ را شناسیم ولی اندازه نمونه خیلی بزرگ باشد.
- (ج) درخصوص ویژگی های تاثیرگذار تردید داشته باشیم.
- (د) ویژگی های تاثیرگذار بر پاسخ را بشناسیم و اندازه نمونه خیلی بزرگ نباشد.

۹۶- کدام یک از ترکیب های زیر جملگی از دلایل پایش (monitoring) در یک کارآزمایی بالینی است؟

- (الف) داده پردازی ، رعایت منشور مطالعه ، تعیین حجم نمونه
- (ب) داده پردازی ، رعایت منشور مطالعه ، اثرات ناسازگار
- (ج) رعایت منشور مطالعه ، اثرات ناسازگار ، تدوین منشور مطالعه
- (د) داده پردازی ، اثرات ناسازگار ، تعیین اهداف مطالعه

۹۷- در مورد تخصیص تصادفی افراد به دو گروه A و B ، اگر ارقام تصادفی ۴-۰ را برای ترتیب AB و ارقام تصادفی ۹-۵ را برای ترتیب BA در نظر بگیریم. این روش مربوط به کدام یک از روش های تصادفی سازی است.

- (الف) کمینه سازی
- (ب) طبقه ای
- (ج) بلوک های تبدیل شده تصادفی درون طبقه
- (د) بلوک جایگشتی تصادفی

۹۸- هدف کار آزمایی در مرحله اول چیست؟

- الف) تعیین دوز دارو
ب) بررسی تاثیر دارو
ج) ایمنی سنجی دارویی
د) بررسی تاثیر و دوز دارو

۹۹- برای مطالعه کدام بیماری انجام کار آزمایی متقاطع مناسب نیست؟

- الف) دیابت ب) آسم ج) آنفولانزا د) پرفشاری خون

۱۰۰- کدام عبارت در مورد روش دوسو کور (Double Blind) صحیح است؟

- الف) روش دوسو کور تنها زمانی عملی است که درمان های ماهیتا مشابه و بیمار از نوع درمان آگاهی ندارند و رضایت دهد.
ب) روش دوسو کور تنها زمانی عملی است که درمان ها ماهیتا مشابه و بیمار و درمانگر از نوع درمان آگاهی ندارند.
ج) درمانگر از نوع درمان آگاهی ندارد ولی بیمار از نوع درمان آگاهی دارد ولی رضایت دهد.
د) بیمار از نوع درمان آگاهی ندارد ولی ملزم به رعایت دستورات درمانگر است.

موفق باشید