بسم الله الرحمن الرحیم

بانک سوالات ده‌سالانه

آمـــارزیستـــی

«کارشناسی ارشد»

(همراه با پاسخنامه تشریحی)

ویژه‌ی رشته‌های: آمارزیستی، اپیدمیولوژی و سایر رشته‌های علوم پزشکی

مولفین و گردآورندگان:

دکتـــر الهام شعـــرباف عیدگاهی

«دکتری تخصصی آمارزیستی، دانشگاه علوم پزشکی شهید بهشتی»

سمیـــه برزنونـــی

«کارشناسی ارشد آمارزیستی، معاونت آموزش، تحقیقات و فناوری دانشگاه علوم پزشکی تربت حیدریه»

دکتـــر اعظـــم صفـــار

«دکتری تخصصی آمارزیستی، گروه آمارزیستی، دانشکده علوم‌پزشکی دانشگاه تربیت مدرس»

طلیعه سخن مؤلف:

در رشته آمارزیستی و در مقطع کارشناسی ارشد، علی‌رغم تمایل دانشجویان بسیاری از رشته‌های متفاوت علوم پزشکی برای شرکت در این رشته، کتاب کامل و با پاسخ تشریحی سوالات کنکور کارشناسی ارشد وزارت بهداشت در آمارزیستی به چاپ نرسیده است. هدف اصلی این مجموعه، آشنا نمودن دانشجویان رشته‌های مختلف وزارت بهداشت با سوالات دوره کارشناسی ارشد آمارزیستی است. علاوه بر این گردآوری و توضیح نکات و مباحث مهم در قالب پاسخنامه تشریحی سوالات کنکور؛ مورد بررسی قرار گرفته است.

کتاب حاضر، پاسخ تشریحی سوالات کنکور کارشناسی ارشد آمارزیستی وزارت بهداشت از سال ۱۳۹۲ تا ۱۴۰۲ است که شامل سوالات ریاضی عمومی، آمار ریاضی و احتمال، و روش‌های آماری است.

قابل ذکر است که کتب مرجع استفاده شده برای پاسخ‌گویی به سوالات دروس ذکر شده به صورت لیست زیر است:

- حساب دیفرانسیل و انتگرال و هندسه تحلیلی جلد ۱ قسمت ۱ و ۲، مولف: جورج توماس، راس فینی، مترجم: مهدی بهزاد، سیامک کاظمی، علی کافی
- ریاضیات عمومی ۱، فرزین حاجی جمشیدی، دانشگاه آزاد تهران مرکزی، انتشارات پوران پژوهش
- مبانی آمار ریاضی، دکتر احمد پارسیان، دانشگاه صنعتی اصفهان
- آمار و احتمال مقدماتی، دکتر جواد بهبودیان، دانشگاه امام رضا (ع)
- Applied Linear Statistical Models - Michael H. Kutner, Christopher J.Nachtsheim, John Neter, William Li- Fifth Edition
- رگرسیون خطی کاربردی، دکتر عبدالرضا بازرگان لاری، دانشگاه شیراز
- آمار ناپارمتری، دکتر جواد بهبودیان، دانشگاه شیراز
- طرح و تحلیل آزمایش‌ها، داگلاس سی. مونت‌گمری، ترجمه دکتر غلامحسین شاهکار، مرکز نشر دانشگاهی، تهران
- خلاصه مباحث اساسی کارشناسی ارشد آمار (آمار نظری- احتمال و کاربرد)، جلد اول، عبدالرضا محمدی، انتشارات پردازش
- خلاصه مباحث اساسی کارشناسی ارشد آمار (آمار نظری- آمار ریاضی)، جلد دوم، عبدالرضا محمدی، انتشارات پردازش

امید است که کتاب حاضر گامی مثبت در جهت پیشرفت شما عزیزان برداشته باشد و در جهت هرآن‌چه بهتر کردن این کتاب خواهشمندیم نظرات و پیشنهادات خود را با ما در میان بگذارید.

با تشکر

مولفین

elham.shaarbaf@gmail.com
s.barzanouni@gmail.com
azam.saffar66@gmail.com

فهرست مطالب

عنوان **صفحه**

کارشناسی ارشد آمارزیستی سال ۹۳-۹۲
سوالات ۷
پاسخنامه تشریحی ۱۷
کارشناسی ارشد آمارزیستی سال ۹۴-۹۳
سوالات ۳۳
پاسخنامه تشریحی ۴۳
کارشناسی ارشد آمارزیستی سال ۹۵-۹۴
سوالات ۶۲
پاسخنامه تشریحی ۷۲
کارشناسی ارشد آمارزیستی سال ۹۶-۹۵
سوالات ۸۹
پاسخنامه تشریحی ۹۹
کارشناسی ارشد آمارزیستی سال ۹۷-۹۶
سوالات ۱۱۸
پاسخنامه تشریحی ۱۲۷
کارشناسی ارشد آمارزیستی سال ۹۸-۹۷
سوالات ۱۴۵
پاسخنامه تشریحی ۱۵۴
کارشناسی ارشد آمارزیستی سال ۹۹-۹۸
سوالات ۱۷۱
پاسخنامه تشریحی ۱۸۰
کارشناسی ارشد آمارزیستی سال ۴۰۰-۹۹
سوالات ۱۹۶
پاسخنامه تشریحی ۲۰۵
کارشناسی ارشد آمارزیستی سال ۴۰۱-۴۰۰
سوالات ۲۲۱
پاسخنامه تشریحی ۲۳۰
کارشناسی ارشد آمارزیستی سال ۴۰۲-۴۰۱
سوالات ۲۵۱
پاسخنامه تشریحی ۲۶۱
کارشناسی ارشد آمارزیستی سال ۴۰۳-۴۰۲
سوالات ۲۷۹
پاسخنامه تشریحی ۲۸۹

سال

۹۲–۹۳

سوالات

کارشناسی ارشد

آمارزیستی

ریاضی عمومی

۱. دامنه پیوستگی تابع $y = \left|\frac{x \sin x}{x^2 + 2}\right|$ کدام است؟

۱) اعداد حقیقی
۲) اعداد حقیقی مثبت
۳) $[0,1]$
۴) $(0,1)$

۲. اگر $x^5 + 4xy^3 - y^5 = 2$، $\frac{dy}{dx}$ کدام است؟

۱) $\frac{5x^4 + 4x^3}{5x^4 - 12xy^2}$
۲) $\frac{5x^4 + 4y^3}{5y^4 - 12xy^2}$
۳) $\frac{5y^4 + 4x^3}{5x^4 - 12xy^2}$
۴) $\frac{5x^4 + 4y^4}{5y^4 - 12xy^2}$

۳. $\lim_{x \to 0} (\frac{1}{\sin x} - \frac{1}{x})$ کدام است؟

۱) ∞
۲) ۱
۳) صفر
۴) $\frac{1}{2}$

۴. $\lim_{x \to \frac{\pi}{2}} \frac{\tan x}{1 + \tan x}$ برابر است با:

۱) $\frac{2}{3}$
۲) $\frac{4}{5}$
۳) $\frac{1}{6}$
۴) ۱

۵. $\int \sqrt{۲x+۱}\,dx$ برابر است با:

۱) $\frac{۱}{۳}(۲x+۱)^{\frac{۳}{۲}}+C$

۲) $(۲x+۱)^{\frac{۳}{۲}}+C$

۳) $\frac{۲}{۳}(۲x+۱)^{\frac{۳}{۲}}+C$

۴) $\frac{۳}{۲}(۲x+۱)^{\frac{۳}{۲}}+C$

۶. $\int_{۰}^{\frac{\pi}{۴}} \tan^۲ x\,dx$ کدام است؟

۱) $۱-\frac{\pi}{۴}$ ۲) $\frac{\pi}{۴}-۱$ ۳) $\frac{\pi}{۴}$ ۴) ۱

۷. اگر تابع $h(x)$ زوج و به ازای همه xها پیوسته باشد آن‌گاه تابع $h(x)\sin x$ تابعی ..... .

۱) زوج است. ۲) فرد است.

۳) نه زوج و نه فرد است. ۴) تعریف نشده است.

۸. اگر تابع $h(x)$ زوج باشد در این صورت $\int_{-a}^{a} h(x)dx$ برابر است با:

۱) $\int_{-a}^{a} -h(x)dx$

۲) $-۲\int_{۰}^{a} h(x)dx$

۳) $۲\int_{۰}^{a} h(x)dx$

۴) ۰

۹. سرعت جسمی که روی یک خط حرکت می‌کند چنین است: $v(t)=۵\pi\cos\pi t \quad m/sec$ کل مسافتی را که جسم از $t=۰$ تا $t=\frac{۳}{۲}$ می‌پیماید کدام است؟

۱) $\frac{۱۵}{۲}$ ۲) ۳۰ ۳) ۱۵ ۴) ۱۰

۱۰. مساحت ناحیه‌ای که از بالا به سهمی $y=۲-x^۲$، از پایین به خط $y=-x$ محدود است برابر است با:

۱) ۳ ۲) ۶ ۳) $\frac{۳}{۲}$ ۴) $\frac{۹}{۲}$

آمار ریاضی و احتمال

۱۱. اگر $X_۱, X_۲, ..., X_n$ هم توزیع مستقل نرمال استاندارد باشند و $\overline{X}_{n-k}=\frac{۱}{n-k}\sum_{i=k+۱}^{n} X_i$ باشد توزیع $\frac{X_۱}{X_۲}$ کدام است؟

۱) مربع کای با n درجه آزادی

۲) مربع کای با $n-۱$ درجه آزادی

۳) $F_{۱,۱}$

۴) کوشی

۱۲. کم‌ترین تعداد پرتاب یک سکه سالم چند بار باشد تا $P(۰/۴ < \bar{X} < ۰/۶)$ حداقل برابر ۰/۹ باشد؟

۱) ۵۰ ۲) ۱۰۰ ۳) ۲۵۰ ۴) ۳۳۵

۱۳. در صورت وجود تابع مولد گشتاور، $Sup(P(X \geq \alpha))$ کدام است؟

۱) $e^{-at}M_x(t)$ ۲) $e^{at}M_x(t)$

۳) $\frac{e^{-at}}{M_x(t)}$ ۴) $\frac{e^{at}}{M_x(t)}$

۱۴. چگالی احتمال توام دو متغیر تصادفی X و Y به شکل زیر است:

$$f(x,y) = xe^{-x(y+1)} \qquad x > ۰, y > ۰$$

تابع مولد گشتاور XY کدام است؟

۱) $(۱-۲t)^{-۱}$ ۲) $(۱-۲t)^{-۲}$

۳) $(t-۱)^{-۱}$ ۴) $(۱-t)^{-۲}$

۱۵. متغیر تصادفی X دارای توزیع نرمال با میانگین صفر و واریانس ۲ است، $E(X^۴)$ برابر است با:

۱) ۱۸ ۲) ۱۲ ۳) ۱۶ ۴) ۶

۱۶. اگر X دارای تابع چگالی $f(x) = \frac{x}{۱۲}; ۱ < x < a$ باشد، $E(X)$ کدام است؟

۱) $\frac{۱-a^۲}{۳۶}$ ۲) $\frac{a^۲-۱}{۳۶}$ ۳) $\frac{۳۱}{۹}$ ۴) $\frac{۱۲۵}{۳۶}$

۱۷. اگر تابع چگالی توام X و Y به شکل زیر باشد، $E(X|Y)$ کدام است؟

$$f(x,y) = \begin{cases} ۸e^{-۲(x+۲y)} & x > ۰, y > ۰ \\ ۰ & \text{other place} \end{cases}$$

۱) $\frac{۱}{۴}$ ۲) $\frac{۱}{۲}$ ۳) ۴ ۴) ۱

۱۸. اگر $X_۱, X_۲, ..., X_n$ یک نمونه تصادفی از توزیع یکنواخت روی فاصله (۱ و ۰) باشد، امید ریاضی بزرگ‌ترین آماره ترتیبی کدام است؟

۱) $\frac{۱}{n(n+۱)}$ ۲) $\frac{n-۱}{n+۱}$

۳) $\frac{n-۱}{n}$ ۴) $\frac{n}{n+۱}$

۱۹. فرض کنید X یک متغیر تصادفی پیوسته با تابع چگالی احتمال f و تابع توزیع F باشد، $E(e^{F(X)})$ برابر است با:

۱) $e-۱$ ۲) $e^{\frac{۱}{۲}}$ ۳) $\frac{۱}{۲}$ ۴) ۱

۲۰. اگر تابع مولد گشتاور متغیری به صورت $(۰/۷۵ + ۰/۲۵e^t)^۵$ باشد احتمال $P(X \geq ۱)$ برابر است با:

۱) ۱ ۲) $(۰/۷۵)^۵$

۳) $۱-(۰/۷۵)^۵$ ۴) $(۰/۲۵)^۵$

۲۱. تعداد متوسط بیمارانی که روزانه وارد یک اورژانس می‌شوند ۳ بیمار است اگر این اورژانس ظرفیت پذیرش ۶ بیمار را داشته باشد احتمال این‌که در یک روز معین ظرفیت اورژانس تکمیل گردد برابر است با:

۱) صفر ۲) $\frac{۸۰}{۸۱}e^{-۳}$ ۳) ۱ ۴) $\frac{۸۱}{۸۰}e^{-۳}$

۲۲. میانگین تعداد بیمارانی که روزانه وارد یک بخش می‌شوند ۲ بیمار است. تجهیزات این بخش قادر است حداکثر ۴ بیمار را در روز سرویس دهد. احتمال این‌که در یک روز معین این بخش نتواند جواب‌گوی بیمارانی که در صف انتظار قرار دارند باشد کدام است؟

۱) $۷e^{-۲}$ ۲) $۱-۷e^{-۲}$

۳) $\frac{e^{-۲}۲^{۴}}{۴!}$ ۴) $۱-\frac{e^{-۲}۲^{۴}}{۴!}$

۲۳. اگر $P(A)=۰/۴$ ، $P(A\cup B)=۰/۲$ و $P(A\mid B)=۰/۲$ باشد مقدار $P(B)$ کدام است ؟

۱) ۰/۳ ۲) ۰/۲۵ ۳) ۰/۲ ۴) ۰/۳۵

۲۴. احتمال مثبت بودن آزمایش برای فرد A برابر ۰/۷ و برای فرد B برابر ۰/۴ می‌باشد. اگر A یا B حتماً در این آزمایش جواب مثبت داشته باشند احتمال این‌که فقط شخص A دارای جواب مثبت باشد کدام است؟

۱) ۰/۲۸ ۲) ۰/۱ ۳) ۰/۱۲ ۴) ۰/۶

۲۵. از هر گروه ۱۵ نفره، پنج نفر دارای افسردگی خفیف می‌باشند اگر از این جامعه ۶ نفر به تصادف انتخاب گردد، متوسط تعداد افرادی که در این نمونه دارای افسردگی خفیف می‌باشند برابر است با:

۱) ۱ ۲) ۳ ۳) ۵ ۴) ۲

۲۶. اگر احتمال تشخیص یک بیماری توسط آزمایشگاه ۵۰ درصد باشد احتمال این‌که در سه بار آزمایش توسط آزمایشگاه مذکور تشخیص صحیح ارایه گردد کدام است؟

۱) $\frac{۳}{۴}$ ۲) $\frac{۷}{۸}$ ۳) $\frac{۴}{۵}$ ۴) $\frac{۹}{۱۰}$

۲۷. اگر نسبت افراد مبتلا به اضطراب در جامعه ۳۰ درصد باشد و آزمونی افراد مبتلا به اضطراب را با احتمال ۰/۸۰ و افراد سالم را با ۰/۲۰ مضطرب قلمداد کند اگر فردی از این جامعه به تصادف انتخاب و جواب آزمایش وی مثبت باشد احتمال این‌که فرد انتخابی دارای اضطراب باشد چقدر است؟

۱) ۰/۸ ۲) ۰/۲ ۳) ۰/۳۲ ۴) ۰/۶۳

۲۸. اگر ده درصد افراد جامعه‌ای مبتلا به دیابت باشند، احتمال این‌که سومین نفری که به تصادف انتخاب می‌شود اولین فردی باشد که دارای دیابت است کدام می‌باشد؟

۱) ۰/۰۸۱ ۲) ۰/۰۰۱ ۳) ۰/۱ ۴) ۰/۸۱۷

۲۹. تابع چگالی طول عمر یک دارو برحسب ساعت به شکل زیر می‌باشد: مقدار $P(X>۲۰)$ کدام است؟

$$f(x)=\begin{cases}\frac{۱}{x^{۲}} & x>۱۰\\ ۰ & x\le ۱۰\end{cases}$$

۱) ۰/۷۵ ۲) ۱ ۳) $\frac{۱}{۲}$ ۴) ۰/۶

۳۰. تابع چگالی X به صورت $f(x)=\begin{cases} a+bx & 0\le x\le 1 \\ 0 & \text{other place} \end{cases}$ می‌باشد. اگر $E(X)=\frac{3}{5}$ باشد مقادیر a و b به ترتیب کدام می‌باشند؟

۱) ۱ و ۲ ۲) ۰/۴ و ۱/۲ ۳) ۰/۶ و ۰/۴ ۴) ۲ و ۰/۶

روش‌های آماری (طرح آزمایش‌ها- رگرسیون- نمونه‌گیری- آمار ناپارامتری)

۳۱. در آنالیز واریانس دوطرفه متعادل با ۴ تکرار، درجه آزادی A و AB با هم برابر و مساوی ۳ می‌باشد، درجه آزادی خطا چقدر است؟

۱) ۲۴ ۲) ۱۲ ۳) ۶ ۴) ۳

۳۲. در آزمون نرمال بودن صفتی در جامعه‌ای، مشاهدات نمونه‌ای را در ۸ گروه طبقه‌بندی و میانگین و واریانس آن را برآورد نموده‌ایم. درجه آزادی آماره آزمون نیکویی برازش χ^2، حداکثر کدام است؟

۱) ۷ ۲) ۴ ۳) ۵ ۴) ۶

۳۳. در جامعه‌ای صفتی دارای توزیع نرمال با واریانس ۹ می‌باشد. اگر بخواهیم $H_0:\mu=15$ را در مقابل $H_1:\mu=18$ در سطح $\alpha=0/05$ و $\beta=0/05$ آزمون کنیم، حداقل چه تعداد نمونه لازم است؟($Z_{0/975}\approx 2$ و $Z_{0/95}\approx 1/65$)

۱) ۱۱ ۲) ۱۴ ۳) ۴۲ ۴) ۳۳

۳۴. یک نمونه تصادفی ۱۰۰ تایی با میانگین μ و واریانس ۴ در نظر می‌گیریم. اگر $(\bar{x}-1,\bar{x}+1)$ یک فاصله اطمینان برای μ باشد، ضریب اطمینان تقریباً چقدر است؟

۱) ۰/۹۰ ۲) ۰/۹۵ ۳) ۰/۵۰ ۴) ۱

۳۵. در طراحی آزمایشات، بلوک‌بندی به چه منظور به کار برده می‌شود؟

۱) کاهش اریبی واریانس برآورد
۲) حذف اثر متغیر مخدوش‌کننده
۳) کنترل سطح اشتباه نوع اول
۴) افزایش درجه آزادی خطا

۳۶. اگر متغیر تصادفی Z دارای میانگین صفر و واریانس ۴ باشد ضریب تغییرات $Y=3Z+6$ برابر است با:

۱) ۲ ۲) ۱ ۳) ۰/۵ ۴) ۳

۳۷. اگر بهره‌هوشی افراد یک جامعه دارای توزیع نرمال با میانگین و واریانس برابر ۱۰۰ باشد، صدک پنجم توزیع تقریباً برابر است با: ($Z_{0/975}=1/96$ و $Z_{0/95}=1/65$)

۱) ۱۱۹/۶ ۲) ۸۰/۴ ۳) ۱۱۶/۵ ۴) ۸۳/۵

۳۸. در مدل رگرسیونی $Y=\beta_0+\beta_1X_1+\beta_2X_2+\varepsilon$ ضریب تعیین $R^2=0/61$ به دست آمده است. اگر $r_{x_1,x_2}>0/5$ و $r_{x_1,y}=0/6$ باشد، $r_{x_2,y}$ کدام‌یک از موارد زیر می‌تواند باشد؟

۱) ۰/۳ ۲) ۰/۱۵ ۳) ۰/۶ ۴) ۰/۱

۳۹. آزمون آماری بدون پارامتر جایگزین آنالیز واریانس با داده‌های تکراری در حالتی که شرایط انجام آزمون پارامتریک برقرار نیست، کدام است؟

۱) فریدمن
۲) آزمون Q ککران
۳) کروسکال والیس
۴) منتل هنزل

۴۰. در یک رگرسیون خطی $\hat{y}=b_0+b_1x$ با ۱۱ مشاهده، اگر $b_1=2$ و $\sum(x_i-\bar{x})^2=100$ و $S_y^2=80$ باشد، ضریب تعیین تقریباً برابر است با:

۱) ۰/۴ ۲) ۰/۵ ۳) ۰/۶ ۴) ۰/۷

۴۱. در یک رگرسیون خطی با مدل $Y_i = \beta_1 x_i + \varepsilon_i$ که در آن $\varepsilon_i \sim N(0,\sigma^2)$ باشد، در یک نمونه n تایی اگر e_iها باقی‌مانده برآورد خط رگرسیون فوق باشد، کدام گزینه همواره درست است؟

۱) $\sum_{i=1}^{n} e_i = 0$

۲) $\sum_{i=1}^{n} \varepsilon_i = 0$

۳) $\sum_{i=1}^{n} x_i e_i = 0$

۴) $\sum_{i=1}^{n} x_i \varepsilon_i = 0$

۴۲. در رگرسیون خطی نرمال $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$، در چه حالتی برآوردهای هر دو پارامتر β_0 و β_1 دارای دقت بالایی هستند؟

۱) $\bar{x}$ صفر باشد و واریانس x زیاد باشد.

۲) $\bar{x}$ صفر باشد و واریانس x کم باشد.

۳) به $\bar{x}$ بستگی ندارد ولی لزوماً باید واریانس x کم باشد.

۴) به $\bar{x}$ بستگی ندارد ولی لزوماً باید واریانس x زیاد باشد.

۴۳. در آزمون فرضیه اگر در ابتدای مطالعه تصمیم بگیریم خطای نوع اول را از ۰/۰۵ به ۰/۱۰ افزایش دهیم کدام گزینه صحیح است؟

۱) توان آزمون کاهش می‌یابد.

۲) احتمال رد فرض صفر کاهش می‌یابد.

۳) احتمال رد فرض صفر افزایش می‌یابد.

۴) احتمال رد فرض صفر و نیز توان آزمون کاهش می‌یابد.

۴۴. برای آزمون فرض $H_0 : \sigma^2 = 64$ با انتخاب نمونه‌ای تصادفی به حجم ۴۶ فرد، آماره آزمون دارای کدام توزیع است؟

۱) x^2 با ۴۴ درجه آزادی

۲) F با یک و ۴۴ درجه آزادی

۳) F با یک و ۴۵ درجه آزادی

۴) x^2 با ۴۵ درجه آزادی

۴۵. فرض کنید رابطه زیر بین صفت کمکی X و صفت اصلی Y در جامعه برقرار است. $y_i = A + Bx_i + \varepsilon_i$ که ε_iها متغیرهای تصادفی با میانگین صفر و واریانس σ^2 هستند برآورد ماکزیمم درستنمایی A عبارت است از: (B معلوم)

۱) $\bar{Y} - B\bar{X}$

۲) $\bar{X} - B\bar{Y}$

۳) $\bar{Y}$

۴) $\bar{X}$

۴۶. در روش نمونه‌گیری لاهیری که در آن $\bar{X}_N$ میانگین صفت کمکی در جامعه‌ای به حجم N و $i=1,\ldots,N$ و $M = M\alpha x(X_i)$ باشد، احتمال این‌که یک زوج انتخاب شده به انتخاب واحدی از جامعه منجر نشود برابر است با:

۱) $\frac{\bar{X}_N}{M}$

۲) $1 - \frac{M}{\bar{X}_N}$

۳) $\frac{M}{\bar{X}_N}$

۴) $1 - \frac{\bar{X}_N}{M}$

۴۷. در یک رگرسیون خطی چندگانه $y_i = \beta + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_{P-1} X_{i,p-1} + \varepsilon_i$ اگر تعداد مشاهدات برابر تعداد پارامترها باشد آن‌گاه همواره:

۱) $R^2 = 0$ ۲) $R^2 = 1$ ۳) $R^2 < 1$ ۴) $R^2 = 0/5$

۴۸. فرض کنید S^2 واریانس برآوردکننده میانگین در نمونه‌ای به حجم $3n$ از توزیع یکنواخت در فاصله (a,b) باشد، چنانچه جامعه را به سه طبقه با طول یکسان طبقه‌بندی کنیم و از هر طبقه n نمونه انتخاب کنیم، آن‌گاه واریانس برآوردکننده میانگین براساس نمونه طبقه‌ای برابر است با:

۱) $\frac{S^2}{3}$ ۲) $3S^2$ ۳) $\frac{S^2}{9}$ ۴) $9S^2$

۴۹. اگر تعدادی داده مثبت و بیش از یک داشته باشیم کدام یک از روابط زیر در مورد آن‌ها برقرار است. ($\bar{x}$ میانگین حسابی، $\bar{x}_G$ میانگین هندسی و $\bar{x}_H$ میانگین هارمونیک می‌باشد).

۱) $\bar{x}_H > \bar{x}_G > \bar{x}$ ۲) $\bar{x}_G > \bar{x}_H > \bar{x}$
۳) $\bar{x}_G < \bar{x}_H < \bar{x}$ ۴) $\bar{x}_H < \bar{x}_G < \bar{x}$

۵۰. اگر صفتی در جامعه دارای میانگین ۱۲ و واریانس صفر باشد میانه و مد به ترتیب کدامند؟

۱) قابل محاسبه نمی‌باشند. ۲) صفر و صفر
۳) ۱۲ و ۰ ۴) ۱۲ و ۱۲

۵۱. چنانچه n حجم نمونه و N حجم جامعه باشد اریبی برآورد نسبتی $R_n = \frac{\bar{Y}_n}{\bar{X}_n}$ برابر است با:

۱) $\frac{Cov(R_n, \bar{X}_n)}{\bar{X}_N}$ ۲) $\frac{-Cov(R_n, \bar{Y}_n)}{\bar{Y}_N}$
۳) $\frac{Cov(R_n, \bar{Y}_n)}{\bar{Y}_N}$ ۴) $\frac{-Cov(R_n, \bar{X}_n)}{\bar{X}_N}$

۵۲. در شهری می‌خواهیم شیوع ابتلا به تالاسمی را برآورد نماییم. با اجرای نمونه‌گیری تصادفی ساده بدون جایگذاری آن‌قدر به انتخاب واحدهای نمونه ادامه می‌دهیم تا ۱۰ نفر مبتلا به تالاسمی مشخص شوند. اگر تا این مرحله ۸۰ نفر مورد آزمایش قرار گرفته باشد، آن‌گاه برآورد نااریب شیوع تالاسمی برابر است با:

۱) $\frac{10}{80}$ ۲) $\frac{9}{80}$ ۳) $\frac{10}{79}$ ۴) $\frac{9}{79}$

۵۳. اگر در آزمون کروسکال والیس اختلاف بین k گروه در سطح α معنی‌دار باشد، جهت مقایسه دوبه‌دوی گروه‌ها، سطح معنی‌داری برای کل مقایسات کدام است؟

۱) $\frac{\alpha}{2k}$ ۲) $\frac{\alpha}{k}$ ۳) α ۴) α^k

۵۴. واریانس صفتی در جامعه‌ای به حجم ۱۶ برابر ۱۵ می‌باشد. نمونه‌ای تصادفی ساده بدون جایگذاری انتخاب می‌نماییم. کوواریانس دو واحد مشخص نمونه در دو انتخاب متوالی برابر است با:

۱) -1 ۲) $+1$ ۳) صفر ۴) $\frac{4}{15}$

۵۵. جامعه‌ای از N=۴ عنصر تشکیل شده است. نمونه‌ای تصادفی بدون جایگذاری به حجم n=۳ انتخاب کرده‌ایم. احتمال استخراج واحد سوم چقدر است؟

۱) $\frac{۱}{۴}$ ۲) $\frac{۱}{۳}$ ۳) $\frac{۳}{۴}$ ۴) $\frac{۱}{۲}$

۵۶. در رگرسیون $y_i = \beta_۰ + \beta_۱ x_i + \varepsilon_i$ چنانچه خطا دارای توزیع نرمال $\varepsilon_i \sim (۰, \sigma^۲)$ باشد چنانچه نمودار باقی‌مانده را در برابر «ترتیب زمان» رسم کنیم و نمودار حاصل یک نوار غیرافقی شیب‌دار باشد، تحلیل آن به کدام‌یک از صورت‌های ذیل است:

۱) واریانس خطا ثابت نیست.

۲) به دلیل حذف اشتباهی عامل $\beta_۰$ از مدل می‌باشد.

۳) نیاز به یک تبدیل مناسب دارد.

۴) y_i ها توزیع نرمال ندارند.

۵۷. از چهار مدل رگرسیونی ذیل چند مدل ذاتاً خطی هستند؟ (قابل تبدیل به مدل خطی هستند).

$$y = e^{\beta_۰ + \beta_۱ x_۱ + \beta_۲ x_۲} . \varepsilon$$

$$y = e^{\beta_۰ + \frac{\beta_۱}{x}} . \varepsilon$$

$$y = \frac{۱}{\beta_۰ + \beta_۱ x_۱ + \beta_۲ x_۲ + \varepsilon}$$

$$y = \frac{۱}{۱ + e^{\beta_۰ + \beta_۱ x_۱ + \beta_۲ x_۲ + \varepsilon}}$$

۱) صفر ۲) یک

۳) سه ۴) چهار

۵۸. به داده‌های مقابل مدل رگرسیون y روی x را با روش کم‌ترین مربعات برازش نموده‌ایم.

x	۰	۲	۴	۶
y	۲	۴	۶	۱۲

مقدار خطای برازش کدام‌یک از نقاط زیر برابر صفر می‌باشد؟

۱) (۴ و۲) ۲) (۶ و۳) ۳) (۶ و۴) ۴) (۴ و۲)

۵۹. در مدل رگرسیون $y = \beta_۰ + \beta_۱ x + \varepsilon$ که از مبدا مختصات عبور نمی‌کند، براساس نمونه‌ای ۲۰ تایی، اطلاعات زیر حاصل شده است:

$$\bar{x} = ۲ \quad \bar{y} = ۴ \quad S_x^۲ = ۲۵ \quad S_y^۲ = ۱۵۰ \quad SS_e = ۵۰$$

با استفاده از روش حداقل مربعات، برآورد عرض از مبدأ چقدر است؟

۱) ۸ ۲) ۲ ۳) -۲ ۴) -۸

۶۰. در تحلیل رگرسیونی ساده کدام‌یک از گزینه‌های زیر همواره صحیح است؟

۱) اگر مجموع خطاهای مشاهده شده بزرگ‌تر از صفر باشد دلیلی بر غیرخطی بودن مدل است.

۲) تبدیل لگاریتمی متغیر پاسخ، مؤثرترین روش تبدیل برای خطی‌سازی است.

۳) نمودار مانده‌ها انحراف از خطی بودن را مشخص می‌کند.

۴) نرمال نبودن توزیع متغیر مستقل موجب غیرخطی شدن رابطه بین دو متغیر می‌شود.

۶۱. در مدل آنالیز واریانس $y_{ij} = \mu + \alpha_i + e_{ij}$ که در آن α_i اثر گروه iام، e_{ij} خطا و μ_i میانگین گروه iام است. کدام عبارت نادرست می‌باشد؟

۱) $\hat{\alpha}_i - \hat{\alpha}_j$ یک برآورد اریب برای $\mu_i - \mu_j$ است.

۲) $\Sigma \alpha_i = 0$

۳) $e_{ij} \sim N(0,1)$

۴) $\mu_i = \mu + \alpha_i$

۶۲. از بین ۱۰۰ کودک ۲۰ نفر به تصادف انتخاب و میانگین قد آن‌ها را براساس سن آن‌ها به روش رگرسیونی برآورد نموده‌ایم. فرض کنید ضریب همبستگی بین قد و سن برابر با r=۰/۸ و واریانس نمونه‌ای قدها برابر با ۵ باشد. در این صورت برآورد واریانس برآورد کننده رگرسیون میانگین قد برابر است با:

۱) ۰/۲۵ ۲) ۰/۰۹ ۳) ۰/۰۷۲ ۴) ۴

۶۳. جامعه‌ای از دو طبقه با اندازه‌های ۱۴ و ۲۱ تشکیل شده است چنانچه بخواهیم نمونه‌ای متناسب با سایز به حجم n=۵ انتخاب نماییم، تعداد نمونه‌های ممکن برابر است با:

۱) $\binom{۳۵}{۵}$ ۲) $\binom{۲۱}{۳} + \binom{۱۴}{۲}$

۳) $۹۱ \times \binom{۱۴}{۲}$ ۴) $۹۱ \times \binom{۲۱}{۳}$

۶۴. در مدل آنالیز واریانس دوراهه که در آن فاکتور اول در r سطح و فاکتور دوم در c سطح آزمایش گردد چنانچه n حجم کل مشاهدات باشد آماره آزمون اثر متقابل دارای چه توزیعی است؟

۱) $F[(n-rc),(c-۱)(r-۱)]$ ۲) $F[(c-۱)(r-۱),(n-rc)]$

۳) $F[(c-۱)(r-۱),(n-۱)(rc-۱)]$ ۴) $F[(n-rc),(n-۱)(rc-۱)]$

۶۵. در یک جدول توافقی سه بعدی (Z, Y, X) که در آن Z متغیر مخدوش کننده است، استقلال X و Y به معنای:

۱) استقلال X و Y در سطوح مختلف Z است.

۲) استقلال X و Z در سطوح مختلف Y است.

۳) استقلال Y و Z در سطوح مختلف X است.

۴) استقلال در بدو X و Y و Z است.

۶۶. اگر در ده مراجعه به بخش اورژانس یک بیمارستان ۳ نفر بستری شده باشند، روش مناسب برای آزمون فرضیه $H_0 : P = ۰/۲$ کدام است؟

۱) آزمون Z ۲) آزمون T

۳) آزمون دوجمله‌ای ۴) آزمون $X^۲$

۶۷. اگر ضریب همبستگی پیرسون بین دو متغیر تصادفی X و Y صفر بوده و توزیع توام دو متغیر نرمال باشد، کدام گزاره زیر مناسب‌تر است؟

۱) X و Y از یکدیگر مستقل هستند. ۲) X و Y از یکدیگر مستقل خطی هستند.

۳) X و Y به هم وابسته هستند. ۴) X و Y رابطه غیرخطی دارند.

۶۸. در یک نمونه ۱۰۲ تایی از نوزادان، همبستگی پیرسون بین قد نوزاد و قد مادر r=۰/۶ به دست آمده است. در این صورت مقدار آماره آزمون برای $H_0 : \rho = 0$ چقدر است؟

۱) ۱۰ ۲) ۷/۵ ۳) ۱ ۴) ۱/۲۵

۶۹. اگر در یک اتوبان بین شهری در یک ماه ۱۲ تصادف رخ داده باشد و فرض گردد که تعداد تصادفات در این اتوبان از توزیع پواسون پیروی می‌کند، مقدار آماره آزمون $H_0 : \lambda = 16$ عبارت است از:

۱) $Z = -1$ ۲) $Z = -\frac{1}{4}$

۳) $Z = 1/25$ ۴) $Z = \frac{1}{4}$

۷۰. برای اندازه‌گیری تکرارپذیری بین دو بررسی برای متغیرهای رتبه‌ای کدام آزمون آماری مناسب می‌باشد؟

۱) کاپا ۲) من‌ویتنی

۳) کروسکال والیس ۴) بارتلت

سال

۹۲-۹۳

پاسخنامه تشریحی

کارشناسی ارشد

آمارزیستی

ریاضی عمومی

۱. گزینه «۱»

$$y = \left| \frac{x \sin x}{x^{۲} + ۲} \right|$$

می‌دانیم تابع $|x|$ در کلیه نقاط دامنه تعریف خود پیوسته است. دامنه توابع کسری نیز برابر است با:

$$D = R - \{\text{ریشه‌های مخرج}\}$$

$$x^{۲} + ۲ = ۰ \Rightarrow x^{۲} = -۲$$

۲. گزینه «۲»

$$x^{۵} + ۴xy^{۳} - y^{۵} = ۲, \frac{dy}{dx} = ?$$

$$\frac{dy}{dx} = \frac{۵x^{۴} + ۴y^{۳}}{۵y^{۴} - ۱۲xy^{۲}}, y' = -\frac{f'_x}{f'_y}$$

۳. گزینه «۳»

$$\lim_{x \to ۰} \left(\frac{۱}{\sin x} - \frac{۱}{x} \right) = ?$$

$$\lim_{x \to ۰} \left[\frac{x - \sin x}{x \sin x} \right] = \frac{۰}{۰} \xrightarrow[H]{\text{با استفاده از قاعده هوپیتال داریم}} \lim_{x \to ۰} \left[\frac{۱ - \cos x}{\sin x + x \cos x} \right] = \frac{۰}{۰} \xrightarrow{H}$$

$$\lim_{x \to ۰} \left(\frac{\sin x}{\cos x + \cos x - x \sin x} \right) = \frac{۰}{۲} = ۰$$

۴. گزینه «۴»

$$\lim_{x\to\frac{\pi}{۲}}\frac{\tan x}{۱+\tan x}=\frac{\infty}{\infty}$$

با استفاده از قاعده هوپیتال داریم:

$$\lim_{x\to\frac{\pi}{۲}}\frac{\tan x}{۱+\tan x}=\lim_{x\to\frac{\pi}{۲}}\frac{\frac{۱}{\cos^{۲}x}}{\frac{۱}{\cos^{۲}x}}=۱$$

یادآوری:

$$\frac{d}{dx}(\tan x)=\sec^{۲}x,\frac{d^{۲}}{dx^{۲}}(\tan x)=۲\sec^{۲}x\tan x,۱+\tan^{۲}x=\sec^{۲}x=\frac{۱}{\cos^{۲}x}$$

۵. گزینه «۱»

$$\int\sqrt{۲x+۱}dx$$

$$U=۲x+۱\Rightarrow du=۲dx\Rightarrow\frac{du}{۲}=dx$$

بنابراین:

$$\int\sqrt{۲x+۱}dx=\frac{۱}{۲}\int\sqrt{u}\ du=\frac{۱}{۲}\int u^{\frac{۱}{۲}}du$$

$$=\frac{۱}{۲}\left(\frac{۲}{۳}u^{\frac{۳}{۲}}\right)=\frac{۱}{۳}u^{\frac{۳}{۲}}=\frac{۱}{۳}(۲x+۱)^{\frac{۳}{۲}}+c$$

۶. گزینه «۱»

$$\int_{۰}^{\frac{\pi}{۴}}\tan^{۲}xdx=\int_{۰}^{\frac{\pi}{۴}}(\sec^{۲}x-۱)dx=\int_{۰}^{\frac{\pi}{۴}}\sec^{۲}x\,dx-\int_{۰}^{\frac{\pi}{۴}}dx$$

$$=\tan x\Big|_{۰}^{\frac{\pi}{۴}}-x\Big|_{۰}^{\frac{\pi}{۴}}=\left[\tan\frac{\pi}{۴}-\tan۰\right]-\left[\frac{\pi}{۴}-۰\right]=[۱-۰]-\frac{\pi}{۴}=۱-\frac{\pi}{۴}$$

۷. گزینه «۲»

می‌دانیم که:

$$f_e.f_o=f_o\Leftarrow\begin{cases}\text{تابع فرد}\leftarrow f_o\\ \text{تابع زوج}\leftarrow f_e\end{cases}$$

بنابراین داریم:

$h(x)$: پیوسته و تابع زوج

$\underbrace{h(x)}_{\text{تابع زوج}}\underbrace{\sin x}_{\text{تابع فرد}}$: تابع فرد است.

۸. گزینه «۳»

$$\int_{-a}^{a} \underbrace{h(x)}_{\text{تابع زوج}} dx = ۲\int_{۰}^{a} h(x)dx$$

$$\int_{-a}^{a} \underbrace{h(x)}_{\text{تابع فرد}} dx = ۰$$

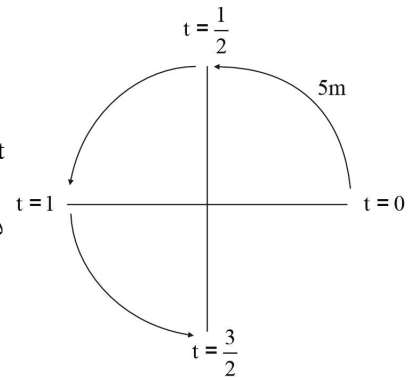
۹. گزینه «۳»

$$f'(t) = v(t) = ۵\pi \cos \pi x \ {}^{m}/_{s}$$

$$f(t) = \int ۵\pi \cos \pi x \, dt = ۵\pi \int \cos \pi x \, dt = \frac{۵\pi}{\pi} \sin \pi t = ۵ \sin \pi t$$

$$t = ۰ \rightarrow x = ۰, t = \frac{۱}{۲} \rightarrow x = ۵, t = ۱ \rightarrow x = ۰, t = \frac{۳}{۲} \rightarrow x = -۵$$

$$۳ \times ۵ = ۱۵$$

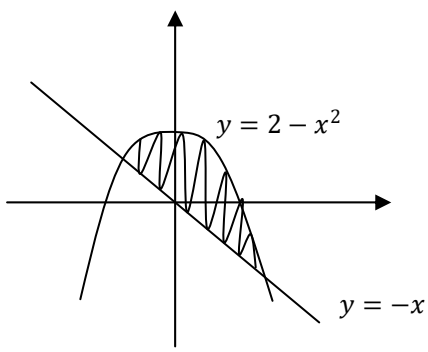


۱۰. گزینه «۴»

$$y = ۲ - x^{۲}, y = -x$$

x	۰	۱	۲
y	۲	۱	-۲

محل تقاطع دو منحنی، نقاط $(۲,-۲)$ و $(۱,+۱)$ است. بنابراین داریم:



$$A = \left| \int_{-۱}^{۲} [(۲ - x^{۲}) - (-x)] dx \right| = \left| \int_{-۱}^{۲} [(۲ - x^{۲} + x)] dx \right|$$

$$= \left| ۲x - \frac{x^{۳}}{۳} + \frac{x^{۲}}{۲} \Big|_{-۱}^{۲} \right| = \left| \left[\left(۴ - \frac{۸}{۳} + \frac{۴}{۲} \right) - \left(-۲ + \frac{۱}{۳} + \frac{۱}{۲} \right) \right] \right| = \frac{۹}{۲}$$

آمار ریاضی و احتمال

۱۱. گزینه «۴»

براساس روابط بین توزیع‌ها اگر $Z_1,...,Z_n$ یک نمونه تصادفی n‌تایی از توزیع $N(0,1)$ باشند، آن‌گاه $U=\frac{z_1}{z_2}\sim C(0,1)$ هم‌چنین $V=\frac{z_1}{|z_2|}\sim C(0,1)$.

۱۲. گزینه «۳»

با استفاده از نامساوی چبیشف می‌دانیم:

$$P(|X-E(X)|<K\sigma)>1-\frac{1}{K^2}$$

با توجه به این نامساوی داریم:

$$P(|X-E(X)|<0/1)>0/9\Rightarrow 1-\frac{1}{K^2}=0/9\Rightarrow\frac{1}{K^2}=0/1\Rightarrow K^2=10$$

از طرفی چون یک سکه سالم داریم، دارای توزیع $B\left(n,\frac{1}{2}\right)$ است. پس داریم:

$$\sigma^2=np(1-p)=n\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)=\frac{n}{4}$$

$$\kappa^2\sigma^2=0/01\Rightarrow 10\times\frac{n}{4}=\frac{1}{100}\Rightarrow n=250$$

۱۳. گزینه «۱»

براساس تعریف sup، می‌دانیم که sup کوچک‌ترین کران بالا می‌باشد. لذا براساس نامساوی مارکوف داریم:

$$P(X\geq a)\leq\frac{E(X)}{a}$$

از طرف دیگر اگر $Y=aX+b$، آن‌گاه $M_Y(t)=e^{bt}M_X(at)$ در این‌جا $Y=a^{-1}Y$ و خواهیم داشت:

$$M_Y(t)=M_X(a^{-1}t)=e^{-at}M_X(t)$$

۱۴. گزینه «۳»

برای رسیدن به تابع مولد گشتاور XY داریم:

$$E(e^{tXY})=\int_0^\infty\int_0^\infty xe^{txy}e^{-xy-x}dydx=\int_0^\infty xe^{-x}dx\int_0^\infty e^{xy(t-1)}dy$$

$$=\int_0^\infty xe^{-x}dx\left[\frac{1}{x(t-1)}e^{xy(t-1)}\Big|_0^\infty\right]=\frac{1}{(t-1)}\int_0^\infty e^{-x}dx(\infty-1)$$

$$=\frac{1}{(t-1)}e^{-x}\Big|_0^\infty=\frac{1}{(t-1)}=(t-1)^{-1}$$

۱۵. گزینه «۲»

$$X \sim N(۰,۲) \qquad E(X^۴) = ? \Rightarrow Z = \frac{X}{\sqrt{۲}} \sim N(۰,۱) \Rightarrow$$

$$Y = Z^۲ = \frac{X^۲}{۲} \sim \chi^۲_{(۱)} \Rightarrow X^۲ = ۲Y \Rightarrow X^۴ = ۴Y^۲$$

$$E(X^۴) = E(۴Y^۲) = ۴E(Y^۲) = ۴\Big[Var(Y) + E^۲(Y) \Big] = ۴(۲ + ۱^۲) = ۱۲$$

یادآوری:

$$X \sim \chi^۲_{(n)} \Rightarrow E(X) = n, Var(X) = ۲n$$

۱۶. گزینه «۳»

$$f(x) = \frac{x}{۱۲}; ۱ < x < a \ , \ E(X) = ?$$

$$E(X) = \int_a^b f(x)dx$$

$$E(X) = \int_۱^a x \frac{x}{۱۲} dx = \frac{۱}{۱۲}\int_۱^a x^۲ dx = \frac{۱}{۱۲}\left[\frac{x^۳}{۳} \Big| a \right]$$

$$= \frac{۱}{۱۲}\left[\frac{a^۳}{۳} - \frac{۱}{۳} \right] = \frac{a^۳}{۳۶} - \frac{۱}{۳۶} = \frac{a^۳ - ۱}{۳۶}$$

از طرفی می‌دانیم که:

$$f(x) = \int_۱^a \frac{x}{۱۲} dx = ۱ \Rightarrow$$

$$\frac{x^۲}{۲۴}\Big|_۱^a = ۱ \Rightarrow \frac{a^۲}{۲۴} - \frac{۱}{۲۴} = ۱$$

$$\Rightarrow \frac{a^۲ - ۱}{۲۴} = ۱ \Rightarrow a^۲ = ۲۵ \Rightarrow a = \pm ۵ \Rightarrow a = ۵$$

در نتیجه داریم:

$$E(X) = \frac{a^۳ - ۱}{۳۶} \xrightarrow{a=۵} E(X) = \frac{۵^۳ - ۱}{۳۶} = \frac{۱۲۴}{۳۶} = \frac{۳۱}{۹}$$

۱۷. گزینه «۲»

$$f(x,y) = \begin{cases} ۸e^{-۲(x+۲y)} & x > ۰, y > ۰ \\ ۰ & \text{other place} \end{cases} \quad , \quad E(X|Y) = ?$$

$$f(x,y) = ۸e^{-۲(x+۲y)} I_{(۰,\infty)}(x) I_{(۰,\infty)}(y) = \Big[۲e^{-۲x} I_{(۰,\infty)}(x) \Big] \Big[۴e^{-۴y} I_{(۰,\infty)}(y) \Big]$$

بنابراین متغیرهای X و Y از هم مستقل می‌باشند. پس:

$$f_X(x) = ۲e^{-۲x} I_{(۰,\infty)}(x) \Rightarrow X \sim E(۲)$$

$$f_Y(y) = ۴e^{-۴y} I_{(۰,\infty)}(y) \Rightarrow Y \sim E(۴)$$

$$f_{X|Y}(x|y) = f_x(x) \Rightarrow E(X \mid Y) = E(X) = \frac{۱}{۲}$$

۱۸. گزینه «۴»

نکته: فرض کنید $X_۱,...,X_n$ یک نمونه تصادفی nتایی از توزیع $U(۰,۱)$ باشند. آنگاه با فرض این‌که $X_{(i)}$، iامین آماره ترتیبی $(i = ۱,...,n)$ باشد، داریم:

$$X_{(i)} \sim Beta(i, n-i+۱)$$

بنابراین توزیع بزرگ‌ترین آماره ترتیبی $(X_{(n)})$ بتا با پارامترهای n و 1 خواهد بود.

$$X_{(n)} \sim Beta(n, ۱)$$

از طرفی می‌دانیم که:

$$X \sim Beta(\alpha,\beta) \Rightarrow E(X) = \frac{\alpha}{\alpha+\beta}$$

در نتیجه خواهیم داشت:

$$E(X_{(n)}) = \frac{n}{n+۱}$$

۱۹. گزینه «۱»

روش اول:

$$E(e^{F(x)}) = \int_{-\infty}^{+\infty} e^{F(x)} f(x)dx = e^{F(x)}\Big|_{-\infty}^{+\infty} = e^{F(+\infty)} - e^{F(-\infty)} = e - ۱$$

روش دوم:

نکته: اگر X متغیر تصادفی پیوسته با تابع توزیع تجمعی F باشد، متغیر تصادفی Y را به صورت $Y = F(X)$ تعریف می‌کنیم. آن‌گاه Y دارای توزیع یکنواخت روی فاصله $(۰,۱)$ است. یعنی:

$$Y = F(X) \sim U(۰,۱)$$

با توجه به این نکته داریم:

$$Y \sim F(X) \sim U(۰,۱), X \sim U(a,b)$$

$$M_Y(t) = \frac{e^t - e^۰}{(۱-۰)t} = \frac{e^t - ۱}{t}, M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}; t \neq ۰$$

$$\Rightarrow E\left[e^{F(x)}\right] = M_Y(t)\Big|_{t=۱} = e - ۱$$

یادآوری:

تابع مولد گشتاور متغیر تصادفی پیوسته X:

$$M_X(t) = E(e^{tx}) = \int e^{tx} f(x)dx$$

۲۰. گزینه «۳»

تابع مولد گشتاور داده شده همان تابع مولد گشتاور توزیع دوجمله‌ای می‌باشد.

$$X \sim Bin(n,p), M_X(t) = (q + pe^t)^n; q = p - ۱$$

بنابراین داریم:

$$M_X(t) = (۰/۲۵e^t + ۰/۷۵)^۵; p = ۰/۲۵,\ n = ۵ \Rightarrow X \sim Bin(۵, ۰/۲۵)$$

$$\Rightarrow P(X \geq ۱) = ۱ - P(X < ۱) = ۱ - P(X = ۰)$$

$$= ۱ - \binom{۵}{۰}(۰/۲۵)^۰(۱ - ۰/۲۵)^{۵-۰} = ۱ - (۰/۷۵)^۵$$

۲۱. گزینه «۴»

$$X \sim P(\lambda = ۳), E(X) = ۳ = \lambda$$

$$X \sim P(\lambda), f_x(x) = \frac{e^{-\lambda}\lambda^x}{x!}; x = ۰,۱,..., \lambda > ۰$$

$$P(X = ۶) = \frac{e^{-۳}۳^۶}{۶!} = \frac{e^{-۳}۷۲۹}{۷۲۰} = e^{-۳}\frac{۸۱}{۸۰}$$

۲۲. گزینه «۲»

$$X \sim P(\lambda = ۲), E(X) = \lambda = ۲$$

$$P(X > ۴) = ۱ - P(X \leq ۴), f(x) = \frac{e^{-\lambda}\lambda^x}{x!}\ x = ۰,...,\infty$$

$$P(X > ۴) = ۱ - P(X \leq ۴) = ۱ - \left[P(X = ۰) + P(X = ۱) + P(X = ۲) + P(X = ۳) + P(X = ۴)\right]$$

$$= ۱ - \left[\frac{e^{-۲}۲^۰}{۰!} + \frac{e^{-۲}۲^۱}{۱!} + \frac{e^{-۲}۲^۲}{۲!} + \frac{e^{-۲}۲^۳}{۳!} + \frac{e^{-۲}۲^۴}{۴!}\right]$$

$$= ۱ - e^{-۲}\left[\frac{۳}{۳} + \frac{۶}{۳} + \frac{۶}{۳} + \frac{۴}{۳} + \frac{۲}{۳}\right] = ۱ - \frac{۲۱}{۳}e^{-۲} = ۱ - ۷e^{-۲}$$

۲۳. گزینه «۲»

$$P(A) = ۰/۲, P(A \cup B) = ۰/۴, P(A|B) = ۰/۲$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(B)P(A|B) = P(B) \times ۰/۲$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = ۰/۲ + P(B) - P(B) \times ۰/۲$$

$$= ۰/۲ + P(B)\left[۱ - ۰/۲\right] = ۰/۲ + ۰/۸P(B) = ۰/۴ \Rightarrow P(B) = \frac{۰/۲}{۰/۸} = ۰/۲۵$$

۲۴. گزینه «۴»

احتمال این‌که فقط شخص A دارای جواب مثبت باشد، برابر است با:

$$P(A^+) = P(A - B) = P(A) - P(A \cap B)$$

از طرفی رابطه زیر برقرار است:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

چون A یا B حتما جواب مثبت دارند، لذا $P(A \cup B) = ۱$ و داریم:

$$-۱ + ۰/۷ + ۰/۴ = P(A \cap B) \Rightarrow P(A \cap B) = ۰/۱ \Rightarrow P(A - B) = ۰/۷ - ۰/۱ = ۰/۶$$

۲۵. گزینه «۴»

احتمال این‌که افراد دارای افسردگی خفیف باشند، برابر است با $p = \frac{n}{N} = \frac{۵}{۱۵}$. در نتیجه می‌توان محاسبه نمود که در یک نمونه ۶ نفری، $np = ۶ \times \frac{۱}{۳}$، ۲ نفر به صورت متوسط دارای افسردگی خفیف هستند.

۲۶. گزینه «۲»

$$X \sim Bin(n = ۳, p = \frac{۱}{۲})$$

$$= P \text{ (هر سه بار تشخیص صحیح باشد)}$$

$$= ۱ - p \text{ (هیچ تشخیص صحیح باشد)} = ۱ - p \text{ (هر سه تشخیص صحیح نباشد)}$$

$$= ۱ - p(X = ۰)$$

$$= ۱ - \binom{۳}{۰}\left(\frac{۱}{۲}\right)^{۰}\left(\frac{۱}{۲}\right)^{۳}$$

$$= ۱ - \frac{۱}{۸} = \frac{۷}{۸}$$

۲۷. گزینه «۴»

اگر A معرف مثبت بودن آزمایش اضطراب و B معرف ابتلا به اضطراب باشد، براساس فرمول بیز داریم:

$$P(B) = ۰/۳, P(\overline{B}) = ۰/۷, P(A|B) = ۰/۸, P(A|\overline{B}) = ۰/۲$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})} = \frac{۰/۸ \times ۰/۳}{۰/۸ \times ۰/۳ + ۰/۲ \times ۰/۷} = \frac{۰/۲۴}{۰/۲۴ + ۰/۱۴} = \frac{۰/۲۴}{۰/۳۸} = ۰/۶۳$$

۲۸. گزینه «۱»

در این‌جا یک توزیع برنولی با $p = ۰/۱$ داریم و نحوه یا روش به دست آوردن احتمال این‌که سومین نفر اولین فردی باشد که دارای دیابت است، به صورت زیر است:

$$(۰/۱)^{۱}(۰/۹)^{۲} = ۰/۰۸۱$$

۲۹. گزینه «۳»

برای محاسبه این احتمال داریم:

$$P(X > ۲۰) = \int_{۲۰}^{\infty} \frac{۱۰}{x^{۲}} dx = -\frac{۱۰}{x}\Big|_{۲۰}^{\infty} = ۰ + \frac{۱۰}{۲۰} = \frac{۱}{۲}$$

۳۰. گزینه «۲»

برای به دست آوردن میانگین داریم:

$$E(X) = \int_{۰}^{۱} (ax + bx^{۲}) dx = a\frac{x^{۲}}{۲} + b\frac{x^{۳}}{۳}\Big|_{۰}^{۱} = \frac{a}{۲} + \frac{b}{۳} = \frac{۳a + ۲b}{۶} = \frac{۳}{۵} \Rightarrow ۱۵a + ۱۰b = ۱۸$$

با توجه به معادله به دست آمده می‌توان $a = ۰/۴$ و $b = ۱/۲$ را در نظر گرفت.

روش‌های آماری (طرح آزمایش‌ها- رگرسیون- نمونه‌گیری- آمار ناپارامتری)

۳۱. گزینه «۱»

درجه آزادی A برابر است با: $a = ۴ \Leftarrow a - ۱ = ۳$.

و درجه آزادی AB نیز از طریق فرمول زیر به دست می‌آید:

$$(a-۱)(b-۱) = ۳(b-۱) = ۳ \Rightarrow b = ۲$$

و چون در آنالیز واریانس متعادل با تکرار، درجه آزادی خطا برابر است با $ab(m-۱)$، که m نشان‌دهنده تکرار است. لذا داریم:

$$df_{\text{خطا}} = ۴ \times ۲ \times (۴-۱) = ۲۴$$

۳۲. گزینه «۳»

وقتی از آزمون $\chi^۲$ جهت بررسی نیکوئی برازش استفاده می‌شود، درجه آزادی آن برابر است با $k-m-۱$، که k تعداد طبقات است و m تعداد پارامترهای برآوردشده. لذا داریم: $k-m-۱ = ۸-۲-۱ = ۵$

۳۳. پاسخ صحیح در گزینه‌ها نیست.

در این سوال آزمون t یک نمونه‌ای را داریم و حداقل حجم نمونه از طریق فرمول زیر به دست می‌آید:

$$n = \frac{\left(Z_{۱-\frac{\alpha}{۲}} + Z_{۱-\beta}\right)^۲ S^۲}{\underbrace{(\mu_۱ - \mu_۰)^۲}_{d}} + \frac{Z^۲_{۱-\frac{\alpha}{۲}}}{۲} = \frac{(۲+۱/۶۵)^۲ \times ۹}{(۳)^۲} + \frac{(۲)^۲}{۲} = ۱۳/۳۲ + ۲ \simeq ۱۶$$

۳۴. گزینه «۴»

فاصله اطمینان برای میانگین به صورت زیر است:

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} Z_{۱-\frac{\alpha}{۲}}, \bar{x} + \frac{\sigma}{\sqrt{n}} Z_{۱-\frac{\alpha}{۲}}\right)$$

$$(\bar{x}-۱, \bar{x}+۱) \Rightarrow \frac{\sigma}{\sqrt{n}} Z_{۱-\frac{\alpha}{۲}} = ۱ \Rightarrow \frac{\sqrt{۴}}{\sqrt{۱۰۰}} Z_{۱-\frac{\alpha}{۲}} = ۱ \Rightarrow Z_{۱-\frac{\alpha}{۲}} = ۵$$

در نتیجه ضریب اطمینان باید برابر با ۱ باشد.

۳۵. گزینه «۲»

شیوه بلوک‌بندی عبارت است از انتخاب واحدها در بلوک‌های متفاوت به‌طوری که واحدهای درون هر بلوک همگن و واحدهای بلوک‌های مختلف متفاوت باشد. این ویژگی منجر به حذف اثر متغیر مخدوشگر می‌شود.

۳۶. گزینه «۲»

ضریب تغییرات Y به صورت $CV_Y = \frac{\sigma_y}{\mu_y}$ محاسبه می‌شود. لذا داریم:

$$E(Y) = ۳E(Z) + ۶ = ۳(۰) + ۶ = ۶, Var(Y) = ۹Var(Z) = ۹ \times ۴ = ۳۶ \Rightarrow CV_Y = \frac{۶}{۶} = ۱$$

۳۷. گزینه «۴»

مقدار صدک برابر است با $\mu + z\sigma$، که z عدد جدول متناظر با مقدار درصد است.

$$z_{۰/۰۵} = -z_{۰/۹۵} = -۱/۶۵$$

$$۱۰۰ - ۱/۶۵ \times ۱۰ = ۱۰۰ - ۱۶/۵ = ۸۳/۵$$

۳۸. گزینه «۳»

می‌دانیم رابطه $R = \frac{r_{y_۲} - r_{۱۲} r_{y_۱}}{\sqrt{۱-r_{۱۲}^۲} - \sqrt{۱-r_{y_۱}^۲}}$ برقرار است، براساس اطلاعات سوال داریم:

چون $r_{۱۲} > ۰/۵$ در نظر گرفته شده است با قرار دادن $r_{۱۲} = ۰/۶$ داریم:

$$\sqrt{۰/۱۶} = \frac{r_{y_۲} - (۰/۶)(۰/۶)}{\sqrt{۱-۰/۳۶}\sqrt{۱-۰/۳۶}} = \frac{r_{y_۲} - ۰/۳۶}{۰/۶۴} \Rightarrow r_{y_۲} = \sqrt{۰/۶۱} \times ۰/۶۴ + ۰/۳۶ = ۰/۸۵$$

چنانچه $r_{۱۲} = ۰/۹$ در نظر گرفته شود:

$$\sqrt{۰/۱۶} = \frac{r_{y_۲} - (۰/۹)(۰/۶)}{\sqrt{۱-۰/۸۱}\sqrt{۱-۰/۳۶}} = \frac{r_{y_۲} - ۰/۵۴}{۰/۳۵} \Rightarrow r_{y_۲} = \sqrt{۰/۶۱} \times ۰/۳۵ + ۰/۵۴ = ۰/۸۱$$

می‌توان به نتیجه رسید که با در نظر گرفتن مقدار تقریبی برای $r_{۱۲}$، $r_{y_۲}$ می‌تواند مقدار ۰/۶ را اختیار کند.

۳۹. گزینه «۱»

آزمون فریدمن برای طرح‌های درون گروهی (نمونه‌های وابسته) مناسب است. این آزمون معادل ناپارامتری آزمون تحلیل واریانس با اندازه‌های تکراری است و تعمیم آزمون ناپارامتری ویلکاکسون می‌باشد.

۴۰. گزینه «۲»

ضریب تعیین برابر است با $R^۲ = \frac{SSR}{SST}$ و با استفاده از روابط در یک مدل رگرسیون داریم:

$$SST = \sum y_i^۲ - \frac{(\sum y_i)^۲}{n} = S_{yy}$$

$$SSR = b_۱^۲ \left(\sum (x_i - \bar{x})^۲ \right) = b_۱^۲ S_{xx}$$

لذا داریم:

$$SSR = ۴ \times ۱۰۰ = ۴۰۰, S_y^۲ = \frac{\sum (y_i - \bar{y})^۲}{n-۱}$$

$$\Rightarrow \sum (y_i - \bar{y})^۲ = S_{yy} = (n-۱) S_y^۲ = ۱۰ \times ۸۰ = ۸۰۰$$

$$R^۲ = \frac{۴۰۰}{۸۰۰} = ۰/۵$$

۴۱. گزینه «۳»

جمله خطا به صورت $\varepsilon_i = y_i - E(y_i)$ تعریف می‌شود که مجهول است و جمله باقی‌مانده بـه صـورت $e_i = y_i - \hat{y}_i$ تعریف می‌شود که معلوم است. از باقی‌مانده‌ها برای تشخیص این‌که آیا مدل رگرسیونی مدل مناسبی است یـا خیـر اسـتفاده می‌شود. از ویژگی‌های باقی‌مانده‌ها داریم:

$$\sum_{i=1}^{n} e_i = 0, \sum_{i=1}^{n} X_i e_i = 0, \sum_{i=1}^{n} \hat{y}_i e_i = 0$$

مجموع مانده‌ها که با مقادیر متغیرهای پیشگو وزن‌دهی شده است یعنی $\sum_{i=1}^{n} X_i e_i$، همیشه برابر با صفر است.

۴۲. گزینه «۱»

خطای استاندارد ضرایب، دقت تخمین ضرایب را اندازه می‌گیرد که مقدار کم‌تر این معیار نشان‌دهنده دقت بیش‌تر است.

$$Var(b_0) = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right), Var(b_1) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}$$

واریانس x برابر است با $S_X^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$، که چون در مخرج کسرهای بالا قرار گرفتـه، اگـر زیـاد شـود منجـر بـه کاهش واریانس ضرایب می‌شود. هم‌چنین $\bar{x} = 0$ نیز باعث بیش‌تر شدن مخرج کسرها و کوچک‌تر شدن واریانس می‌گردد.

۴۳. گزینه «۳»

p- مقـدار کـم‌تـرین مقـداری از α اسـت کـه یافتـه آمـاره آزمـون ممکـن اسـت موجـب رد فـرض H_0 گـردد و اگـر $P - value < \alpha \Rightarrow RH_0$. لذا زمانی که α از ۰/۰۵ به ۰/۱ افزایش پیدا کند، احتمال رد فرض H_0 نیز افزایش می‌یابد.

۴۴. گزینه «۴»

برای آزمون آماری $H_0 : \sigma^2 = 64$ دو حالت وجود دارد: ۱) زمانی که μ معلوم اسـت کـه $\frac{ns^2}{\sigma^2}$ تحـت فـرض H_0 دارای توزیع $\chi^2(n)$ است. ۲) زمانی که μ مجهول است که $\frac{(n-1)s^2}{\sigma^2}$ تحت فرض H_0 دارای توزیع $\chi^2(n-1)$ است. لذا در این‌جا آماره آزمون با در نظر گرفتن مجهول بودن μ دارای توزیع $\chi^2(45)$ است.

۴۵. گزینه «۱»

بـه‌طـور کلـی تـابع چگـالی مشـاهدات Y_i زمـانی کـه جمـلات خطـای مـدل رگرسـیونی نرمـال اسـت و مـدل بـه صورت $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ است، به صورت زیر است:

$$L_i = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)^2\right]$$

تابع درستنمایی براساس ۳ پارامتر β_0، β_1 و σ^2 به صورت زیر است:

$$L(\beta_0,\beta_1,\sigma^2)=\prod_{i=1}^{n}\frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left[-\frac{1}{2\sigma^2}(Y_i-\beta_0-\beta_1X_i)^2\right]$$

$$=\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1X_i)^2\right]$$

با مشتق گرفتن تابع درستنمایی نسبت به هر کدام از پارامترها می‌توان برآورد β_0، β_1 و σ^2 را محاسبه نمود.

$$b_1=\frac{\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sum(X_i-\overline{X})^2},b_0=\frac{1}{n}\left(\sum Y_i-b_1\sum X_i\right)=\overline{Y}-b_1\overline{X}$$

$$\hat{\sigma}^2=\frac{\sum(Y_i-\hat{Y}_i)^2}{n}$$

در نتیجه برآورد درستنمایی A برابر است با $\overline{Y}-B\overline{X}$.

۴۶. گزینه «۴»

در روش نمونه‌گیری لاهیری فرض کنید $\max x_i = M \quad , i=1,...,n$، مراحل به صورت زیر است:

۱) انتخاب جفت اعداد تصادفی (i,j) که $1\le i\le N$ و $1\le j\le M$.

۲) اگر $j\le X_i$، آن‌گاه iامین واحد انتخاب می‌شود و در غیر این صورت جفت اعداد تصادفی دیگر انتخاب می‌شوند.

۳) برای دستیابی به حجم نمونه n، این فرآیند تا زمانی که n واحد انتخاب شوند، تکرار می‌شود.

احتمال انتخاب به صورت $P_c^*=P(1\le i\le N)P(1\le j\le M|i)=\frac{1}{N}.\frac{X_i}{M}$ است و احتمال عدم انتخاب یک واحد به صورت زیر تعریف می‌شود:

$$\frac{1}{N}\sum_{i=1}^{N}\left(1-\frac{X_i}{M}\right)=\frac{1}{N}\left(N-\frac{N\overline{X}}{M}\right)=1-\frac{\overline{X}}{M}$$

۴۷. گزینه «۲»

می‌دانیم که ضریب تعیین با افزایش متغیرهای مستقل افزایش پیدا می‌کند. با توجه به فرمول ضریب تعیین تعدیل شده داریم:

$$R_a^2=1-\frac{n-1}{n-p}.\frac{SSE}{SST}$$

اگر تعداد مشاهدات و تعداد پارامترها برابر باشند، یعنی $n=p$ در نتیجه $R_a^2=1$.

۴۸. گزینه «۳»

اگر $Y\sim U(a,b)$ باشد و واریانس آن به صورت V_1 تعریف شود، آن‌گاه اگر برد متغیر را به L قسمت مساوی تقسیم کنیم واریانس جدید به صورت $V_2=\frac{1}{L^2}\times V_1$ تعریف می‌شود. در این‌جا داریم:

$$Var=\frac{1}{3^2}\times S^2=\frac{S^2}{9}$$

۴۹. گزینه «۴»

میانگین حسابی برابر است با $\bar{x}=\frac{\sum_{i=1}^{k} f_i x_i}{n}$ و میانگین هندسی $\bar{x}_G=\sqrt[n]{X_1^{f_1}X_2^{f_2}...X_k^{f_k}}$ و میانگین هارمونیک $\bar{X}_H=\frac{1}{n}\sum_{i=1}^{k}\frac{f_i}{x_i}$. چون میانگین ریشه‌ای رتبه r یعنی $M_r=\left(\sum_{i=1}^{k}\frac{f_i x_i^{\,r}}{n}\right)^{1/r}$ تابعی صعودی از r است، با قرار دادن r برابر با ۱-، ۰ و ۱ در M_r می‌توان به نامساوی $\bar{x}_H \le \bar{x}_G \le \bar{x}$ دست یافت.

۵۰. گزینه «۴»

چون واریانس برابر با صفر است لذا تمامی مشاهدات با هم برابر است، لذا چون میانگین برابر با ۱۲ است، پس میانه و مد نیز برابر با ۱۲ خواهند بود.

۵۱. گزینه «۴»

اگر برآورد نسبتی به صورت $R_n=\frac{\bar{Y}_n}{\bar{X}_n}$ باشد، برای به دست آوردن اریبی این برآورد داریم:

$$\bar{Y}_n = R_n \times \bar{X}_n$$

$$E(R_n)-\frac{\mu_Y}{\mu_X}=E(R_n)-\frac{E(R_n\bar{X}_n)}{E(\bar{X}_n)}=-\frac{Cov(R_n,\bar{X}_n)}{\bar{X}_N}$$

۵۲. گزینه «۴»

در نمونه‌گیری تصادفی ساده بدون جایگذاری برآورد نااریب شیوع برابر با $\frac{n-1}{N-1}=\frac{9}{79}$ است.

۵۳. گزینه «۳»

در مقایسه‌های جفتی بین k گروه در آزمون کروسکال والیس سطح معنی‌داری برابر با $\frac{\alpha}{2k}$ است. این انتخاب برای سطح معنی‌داری در مقایسه‌های جفتی باعث می‌شود سطح معنی‌داری برای کل مقایسات برابر با α باشد.

۵۴. گزینه «۱»

در نمونه‌گیری تصادفی ساده با جایگذاری مقادیر نمونه از هم مستقلند و کوواریانس بین دو مقدار متفاوت نمونه صفر است، اما هنگامی که نمونه‌گیری تصادفی ساده بدون جایگذاری باشد، کوواریانس بین دو مقدار متفاوت به صورت $Cov(X_i,X_j)=-\frac{\sigma^2}{N-1}; i\neq j$ است. لذا داریم:

$$Cov(X_i,X_j)=-\frac{15}{15}=-1$$

۵۵. گزینه «۳»

از آن‌جا که نمونه‌ها به صورت تصادفی بدون جایگذاری انتخاب می‌شوند، لذا برای انتخاب نمونه ۳ عضوی از جامعه‌ای به حجم ۴، احتمال انتخاب هر ۳ عضو نمونه برابر با $P(A)=\frac{3}{4}$ است.

۵۶. گزینه «۲»

در ترسیم نمودار مانده‌ها در برابر زمان ممکن است ۴ حالت زیر رخ دهد:

۱) مانده‌ها به صورت قیف مانند در طی زمان قرار می‌گیرند. $<$ که نشان‌دهنده این است، واریانس ثابت نیست و به مرور زمان افزایش می‌یابد که باید تجزیه و تحلیل رگرسیون کم‌ترین توان‌های دوم موزون استفاده شود.

۲) یک نوار غیرافقی شیب‌دار که مربوط به خطای تجزیه و تحلیل است و دلیل آن می‌تواند حذف اشتباهی عامل β_0 از مدل باشد.

۳) مانده‌ها به صورت سهمی مانند قرار می‌گیرند. $\cap$ که مدل مناسب نیست و نیاز به جمله یا جملاتی اضافی در مدل است و یا نیاز به یک تبدیل مناسب است.

۴) چنان‌چه نمودار پراکنش یک خوشه بی الگو باشد، فرض وجود استقلال ε_iها را القا می‌کند.

۵۷. گزینه «۴»

منظور از رابطه خطی در مدل رگرسیون، وجود رابطه خطی بین ضرایب است و نه بین متغیرهای مستقل. بر این اساس می‌توان مشاهده نمود هر ۴ مدل دارای رابطه خطی بین ضرایب می‌باشند.

$$y = e^{\beta_0+\beta_1 x_1+\beta_2 x_2}.\varepsilon \Rightarrow \ln y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ln\varepsilon \quad \checkmark$$

$$y = e^{\beta_0+\frac{\beta_1}{x}}.\varepsilon \Rightarrow \ln y = \beta_0 + \beta_1 x^{-1} + \ln\varepsilon \quad \checkmark$$

$$y = \frac{1}{\beta_0+\beta_1 x_1+\beta_2 x_2+\varepsilon} \quad \checkmark \qquad y = \frac{1}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\varepsilon}} \quad \checkmark$$

۵۸. گزینه «۲»

هر چه همبستگی بین مقادیر X و Y ضعیف‌تر باشد، پراکندگی نقاط مختصات بیش‌تر می‌شود و هر چه همبستگی قوی‌تر باشد، نمرات به خط رگرسیون نزدیک‌تر می‌شوند. اگر $r = 1$ باشد، همه نقاط روی خط رگرسیون قرار می‌گیرند. خط رگرسیون در نقطه $(\overline{X}, \overline{Y})$ دارای مقداری خطای برازش برابر با صفر است. لذا داریم: $(\overline{X}, \overline{Y}) = (3, 6)$

۵۹. گزینه «۱»

با استفاده از روابط موجود برای رگرسیون داریم:

$$b_0 = \overline{Y} - b_1\overline{X}, SSR = b_1^2 S_{XX} \Rightarrow b_1 = \sqrt{\frac{SSR}{S_{XX}}}, SST = SSR + SSE$$

$$SST = S_{YY} = (n-1)S_Y^2 = 19\times 150 = 2850, S_{XX} = (n-1)S_X^2 = 19\times 25 = 475$$

$$SSR = SST - SSE = 2850 - 50 = 2800$$

$$b_1 = \sqrt{\frac{2800}{475}} = \begin{cases} b_1 = 2/43 \\ b_1 = -2/43 \end{cases}, b_0 = \begin{cases} 4 - 2/43(2) = -0/86 \\ 4 + 2/43(2) = 8/86 \simeq 8 \end{cases}$$

۶۰. گزینه «۳»

در تحلیل رگرسیون از نمودار مانده‌ها در ۳ جایگاه متفاوت استفاده می‌شود که هر کدام ویژگی خاصی را بررسی می‌کند. نمودار مانده‌ها در برابر ترتیب زمان برای بررسی استقلال مانده‌ها استفاده می‌شود. نمودار مانده‌ها در برابر مقدار پیش‌بینی‌شده $(\widehat{Y}_i)$ برای بررسی یکسان بودن واریانس خطاها به کار می‌رود و نمودار مانده‌ها در برابر متغیر مستقل می‌تواند برای بررسی ثبات واریانس خطا و هم‌چنین بررسی خطی‌بودن استفاده شود و انحراف از خطی بودن را مشخص می‌کند.

۶۱. این سوال حذف شده است.

۶۲. گزینه «۳»

واریانس برآوردکننده رگرسیونی برابر است با:

$$\mathrm{Var}(\overline{Y}_d)=\left(\frac{1}{n}-\frac{1}{N}\right)S_Y^2(1-\rho^2)=\left(\frac{1}{20}-\frac{1}{100}\right)5(1-0/64)=\frac{4}{100}\times 5\times\frac{36}{100}=0/072$$

۶۳. گزینه «۴»

تخصیص متناسب براساس قاعده $\frac{n_h}{n}=\frac{N_h}{N}$ می‌باشد. بر این اساس داریم:

$$\frac{N_h}{N}=\frac{5}{35}=\frac{1}{7},\frac{n_1}{14}=\frac{1}{7}\Rightarrow n_1=2,\frac{n_2}{21}=\frac{1}{7}\Rightarrow n_1=3$$

تعداد نمونه‌های ممکن برابر است با:

$$\binom{14}{2}\binom{21}{3}=91\times\binom{21}{3}$$

۶۴. گزینه «۲»

در آنالیز واریانس دوطرفه درجه آزادی اثر متقابل از حاصلضرب تعداد فاکتور اول − ۱ و تعداد فاکتور دوم − ۱ به دست می‌آید و زمانی که تعداد آزمودنی‌ها را داشته باشیم، تعداد آزمودنی‌ها − تعداد تیمارها می‌شود. لذا درجه آزادی برابر است با:

$$\left[(r-1)(c-1),(n-rc)\right]$$

۶۵. گزینه «۱»

چون در جدول توافقی $Z=(X,Y,Z)$ یک متغیر مخدوشگر است، استقلال X و Y به این مفهوم است که دو متغیر X و Y باید در سطوح مختلف Z مستقل باشند.

۶۶. گزینه «۳»

افراد مراجعه‌کننده به بیمارستان یا بستری شده و یا بستری نمی‌شوند. در نتیجه متغیر دو حالتی داریم که دارای توزیع دوجمله‌ای است که برای nهای بزرگ دارای توزیع $N(0,1)$ است.

۶۷. گزینه «۱»

اگر جفت (X,Y) دارای توزیع نرمال استاندارد باشد و ضریب همبستگی آن‌ها $\rho=0$ باشد، آن‌گاه X و Y مستقل‌اند اگر و فقط اگر $\rho=0$ باشد.

۶۸. گزینه «۲»

با استفاده از تبدیل فیشر برای $n\geq 25$ داریم:

$$W=\frac{1}{2}\ln\frac{1+R}{1-R},\mu_W=\frac{1}{2}\ln\frac{1+\rho}{1-\rho},\sigma_W^2=\frac{1}{n-3}$$

به‌طور تقریبی $Z=\frac{W-\mu_W}{\sigma_W^2}\sim N(0,1)$. اما در حالت خاصی که X و Y هر دو نرمال و مستقل باشند $V=\frac{R\sqrt{n-2}}{\sqrt{1-R^2}}$ دقیقا دارای توزیع $T(n-2)$ است که این آزمون معادل با آزمون $H_0:\rho=0$ است. لذا داریم:

$$V=\frac{0/6\sqrt{102-2}}{\sqrt{1-0/36}}=\frac{6}{0/8}=7/5$$

۶۹. گزینه «۱»

در توزیع پواسن مقدار میانگین و واریانس به صورت زیر محاسبه می‌شود:

$$f(x)=\frac{e^{-\lambda}\lambda^{x}}{x_i};\ x=0,1,\ldots \Rightarrow E(X)=\lambda, Var(X)=\lambda$$

در نتیجه آماره آزمون برای فرض $H_0:\lambda=16$ زمانی که متغیر دارای توزیع پواسن $\lambda=12$ است، برابر است با:

$$E(X)=16, Var(X)=16 \Rightarrow Z=\frac{12-16}{\sqrt{16}}=\frac{-4}{4}=-1$$

۷۰. این سوال حذف شده است.