

کوئیز دوم: رگرسیون و استنتاج آماری

تاریخ: ۳۱ فروردین ۱۳۹۹، ساعت ۱۴ به صورت غیر حضوری

زمان: ۶۰ دقیقه برای پاسخ به سوالات، آماده‌سازی فایل، آپلود و ارسال

نکته: لازم است پاسخ کوئیز را به صورت یک فایل PDF در سامانه درس‌افزار آپلود نمایید. در صورت بروز مشکل و در شرایط اضطراری می‌توانید پاسخ‌نامه را قبل از موعد به آدرس ایمیل درس ارسال نمایید. به ازای هر یک دقیقه تاخیر در ارسال و یا آپلود فایل ۱۰ درصد از نمره شما کسر خواهد شد.

بسیار مهم: عبارت زیر را حتما بالای برگه پاسخ‌نامه بنویسید و امضا کنید.

«این جانب (نام و نام خانوادگی) متعهد می‌شوم در پاسخ به سوالات این کوئیز مانند جلسات حضوری امتحان عمل کنم.»

برای ارزیابی تاثیر اندازه مدرسه روی عملکرد تحصیلی دانش‌آموزان داده‌های ۴۰۸ مدرسه در ایالت میشیگان آمریکا جمع‌آوری شده است. جداول ۱ و ۲ نتایج رگرسیون‌های مختلفی که با استفاده از این اطلاعات تخمین زده شده و آمارهای توصیفی متغیرها را نشان می‌دهند. تعریف متغیرهای استفاده شده در ذیل جدول ۱ آورده شده است.

۱. ابتدا توضیح دهید برای پاسخ به سوال فوق ضرایب چه متغیری (یا متغیرهایی) برای ما مهم هستند؟ از لحاظ تئوری انتظار دارید که علامت ضرایب این متغیر (یا متغیرها) چه باشد؟ آیا ستون (۱) با انتظار شما همخوانی دارد؟ ضریب تخمین‌زده شده در این ستون را تفسیر کنید. (۲۰)

۲. توضیح دهید تفسیر علیّ یک ضریب به چه معنی است. تمام فروض مورد نیاز برای تفسیر علیّ ضریب enroll ستون (۱) را نوشته و توضیح دهید که انتظار دارید هر یک از این فروض صادق باشند یا خیر و چرا؟ (۳۰ نمره)

۳. ضریب enroll تخمین‌زده شده در ستون (۲) را تفسیر کنید. از مقایسه ستون‌های (۱) و (۲) توضیح دهید که مقدار ارباب این ضریب در ستون (۱) چگونه بوده است. آیا می‌توانید بگویید به طور متوسط ویژگی‌های مدارس پرجمعیت‌تر چه تفاوتی با مدارس دیگر داشته است؟ (۲۰ نمره)

۴. تفاوت ستون (۳) با ستون (۲) در چیست؟ تفسیر ضرایب تخمین‌زده شده برای متغیر(های) مورد توجه ما چگونه تغییر می‌کند؟ (۱۵ نمره)

۵. توضیح دهید که معنی داری ضریب enroll در سطح ۵ درصد چگونه ارزیابی می‌شود. آیا این ضریب درستون‌های (۱) تا (۳) معنی‌دار است؟ (۱۵ نمره)

جدول ۱. نتایج رگرسیون‌های مختلف برای بررسی تاثیر اندازه مدرسه روی عملکرد تحصیلی دانش‌آموزان

	(1)	(2)	(3)	(4)
enroll	0.00028 (0.00019)	-0.00020 (0.00022)		
staff		0.04792 (0.03981)		
totcomp		0.00046 (0.00010)		
Ln(enroll)			-1.268 (0.693)	-1.595 (0.700)
Ln(staff)			3.98 (4.190)	46.189 (16.93)
Ln(totcomp)			21.155 (4.056)	62.88 (16.716)
Ln(expend)				-43.132 (16.77)
Constant	23.36 (0.73)	2.27 (6.113)	-207.67 (48.70)	-478.29 (115.81)
Observations	408	408	408	408
R^2	0.005	0.054	0.065	0.081
\bar{R}^2	0.003	0.047	0.058	0.071

نکات: هر ستون جدول نتایج یک رگرسیون را نشان می‌دهد. متغیر وابسته در همه ستون‌ها درصدی از دانش‌آموزان مدرسه است که در امتحان ریاضی پایه دهم قبول شده‌اند. enroll: تعداد دانش‌آموزان ثبت‌نامی مدرسه، staff: تعداد کارکنان مدرسه به ازای ۱۰۰۰ دانش‌آموز، expend: متوسط هزینه‌های مدرسه برای هر دانش‌آموز (دلار) و totcomp: متوسط حقوق و مزایای سالانه معلمان مدرسه (دلار) را نشان می‌دهند. $\ln(y)$ نشانگر لگاریتم طبیعی متغیر y است. اعداد در پرانتز خطای استاندارد (standard error) تخمین ضرایب را نشان می‌دهد.

جدول ۲. آمارهای توصیفی متغیرهای استفاده شده در رگرسیون‌های جدول ۱

متغیر	مشاهدات	متوسط	انحراف معیار	مینیمم	ماکزیمم
math10	408	24.11	10.49	1.9	66.7
enroll	408	2663.81	2696.8	212	16793
staff	408	100.6	13.30	65.9	166.6
totcomp	408	38237.9	5985.09	24498	63518
expend	408	4376.58	775.79	3332	7419

نکات: تعریف متغیرها مانند جدول ۱ است. متغیر math10 درصد دانش‌آموزانی که در آزمون ریاضی پایه دهم قبول شده‌اند را نشان می‌دهد.