

میان ترم: ابتدای ترم تا انتهای مبحث متغیر مجازی

تاریخ: ۸ خرداد ۱۳۹۹، ساعت ۹ صبح، غیر حضوری، مجموعاً ۱۰۰ نمره

زمان: ۱۵۰ دقیقه برای پاسخ به سوالات، آماده سازی فایل، آپلود و ارسال

نکته: لازم است پاسخ کوئیز را به صورت یک فایل PDF در سامانه درس افزار آپلود نمایید. در صورت بروز مشکل و در شرایط اضطراری می توانید پاسخ نامه را قبل از موعد از طریق ایمیل ارسال نمایید. به ازای هر یک دقیقه تاخیر در ارسال پاسخ ۵ درصد نمره شما کسر خواهد شد.

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

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

وزن نوزاد هنگام تولد یکی از شاخص های مهمی است که در مطالعات اقتصادی به عنوان معیاری از سلامتی نوزاد استفاده می شود. (در این سوال کاهش وزن را معادل کاهش سلامتی نوزاد در نظر بگیرید.) جدول شماره یک نتایج رگرسیون های مختلفی را نشان می دهد که با استفاده از اطلاعات حدود ۱۸۰۰ نوزاد به دست آمده اند. هم چنین جدول شماره دو خلاصه آماری متغیرهای استفاده شده در این سوالات را نشان می دهد. تعریف متغیرها در توضیحات زیر جدول یک قابل مشاهده است. با توجه به این دو جدول به سوالات زیر پاسخ دهید:

۱. تفسیر ضریب متغیر npvis در ستون (۱) را بیان کنید. آیا این ضریب معنی دار است؟ آیا ضریب npvis در این

ستون با انتظار شما همخوانی دارد؟ (۱۰ نمره)

۲. آزمون کنید که آیا ضریب npvis در ستون (۱) برابر ۰/۰۰۵ هست یا نه؟ (۵ نمره)

۳. تفاوت بین اثر علی و هم بستگی را بیان کنید. برای تفسیر علی اثر افزایش تعداد ویزیت روی وزن نوزاد در ستون

(۱) چه فرضی باید برقرار باشند؟ آیا آن ها برقرار هستند؟ (۱۰ نمره)

۴. در ستون دوم متغیرهای دیگری به رگرسیون اضافه شده اند. آیا ضرایب این متغیرها (علامت و معنی داری) با

انتظار شما همخوانی دارد؟ (۵ نمره)

۵. چرا ضریب متغیر npvis در ستون دو کاهش یافته است؟ علت آن را به طور کامل توضیح دهید. (۵ نمره)

۶. در ستون سوم، توان دوم متغیر npvis (npvissq) و متغیر mage (magesq) نیز اضافه شده اند. با توجه به

نتایج این ستون و ستون های قبل اثرگذاری تعداد ویزیت ها و سن مادر را روی وزن نوزاد به طور کامل تحلیل

کنید. با توجه به این ستون افزایش تعداد ویزیت‌ها از ۱۱ به ۱۲ به طور متوسط چه میزان روی وزن نوزاد اثرگذار است؟ (۱۰ نمره)

۷. در ستون (۴) متغیر fstmon نیز اضافه شده است. تفسیر ضریب این متغیر چیست؟ آیا این ضریب با انتظار شما همخوانی دارد؟ علت را توضیح دهید. (۱۰ نمره)

۸. در ستون (۵) حاصل ضرب دو متغیر fstmon و npvis نیز اضافه شده است. ضریب این متغیر چه چیزی را نشان می‌دهد؟ توضیح دهید. (۱۰ نمره)

۹. R^2 به چه معنی است؟ تفاوت آن با \bar{R}^2 (یا همان Adjusted R-squared) چیست؟ برای هر یک از رگرسیون‌های ستون‌های (۱) تا (۵) کدام یک از این ضرایب مناسب‌تر هستند؟ آیا کوچک بودن این دو (R^2 و \bar{R}^2) به این معنی نیست که ضرایب تخمین زده شده در این جدول نشان‌دهنده رابطه علی نیستند؟ (۱۰ نمره)

۱۰. آزمون معنی‌داری کلی رگرسیون (overall significance) ستون (۲) را به صورت یک آزمون فرضیه استاندارد بنویسید. آماره این آزمون چه توزیعی خواهد داشت؟ این آزمون به چه معنی است؟ (دقت کنید که نیازی به انجام آزمون در این سوال وجود ندارد). (۱۰ نمره)

۱۱. توضیح دهید p-value یک تست به چه معنی است؟ p-value آزمون معنی‌داری کلی رگرسیون ستون (۲) برابر ۰/۰۰۰ است. معنی آن چیست؟ (۱۰ نمره)

۱۲. آیا نتایج رگرسیون ستون (۵) در مورد اثر مراجعه به پزشک در طول بارداری بر روی وزن نوزاد به لحاظ اقتصادی معنی‌دار است؟ یعنی آیا این مساله اهمیت دارد یا خیر؟ توضیح دهید. (۵ نمره)

جدول ۱: نتایج رگرسیون

VARIABLES	(1)	(2)	(3)	(4)	(5)
npvis	0.00663 (0.00132)	0.00486 (0.00130)	0.0125 (0.00368)	0.0141 (0.00372)	0.0120 (0.00378)
cigs		-0.00313 (0.00115)	-0.00279 (0.00116)	-0.00287 (0.00116)	-0.00292 (0.00115)
meduc		-0.000268 (0.00247)	-0.00103 (0.00248)	-0.000798 (0.00248)	-0.000342 (0.00248)
mage		0.00138 (0.00108)	0.0208 (0.00940)	0.0222 (0.00939)	0.0235 (0.00938)
male		0.0251 (0.00969)	0.0257 (0.00967)	0.0259 (0.00965)	0.0254 (0.00963)
npvissq			-0.000267 (0.000119)	-0.000301 (0.000119)	-0.000306 (0.000119)
agesq			-0.000329 (0.000157)	-0.000350 (0.000157)	-0.000371 (0.000157)
fstmon				-0.0299 (0.0108)	-0.140 (0.0387)
fstmon × npvis					0.00891 (0.00301)
Constant	8.037 (0.0161)	8.015 (0.0404)	7.696 (0.138)	7.665 (0.138)	7.664 (0.138)
Observations	1,764	1,644	1,644	1,644	1,644
Adjusted R-squared	0.014	0.016	0.021	0.024	0.029
R-squared	0.0142	0.0189	0.0247	0.0292	0.0343

نکات: هر ستون در این جدول نتایج یک رگرسیون را نشان می‌دهد. متغیر وابسته در تمام رگرسیون‌ها برابر لگاریتم وزن نوزاد (گرم) است. npvis تعداد ویزیت دکتر در طول بارداری است. cigs تعداد سیگارهایی است که مادر باردار در یک هفته می‌کشد. medu سال‌های تحصیل مادر، mage سن مادر، male متغیر مجازی است که برای نوزاد پسر برابر یک و برای دختر برابر صفر است. npvissq توان دوی npvis و agesq توان دوی mage است. fstmon یک متغیر مجازی است که این صورت که اگر مادر تا انتهای ماه اول بارداری به پزشک مراجعه کرده باشد برابر یک است و اگر نه برابر صفر است. انحراف معیار ضرایب در داخل () در زیر آن‌ها نوشته شده است.

جدول ۲: خلاصه آماری

Variable	Obs	Mean	Std. Dev.	Min	Max
bwght	1,832	3401	577	360	5204
npvis	1,764	11.6	3.7	0	40
cigs	1,722	1.1	4.2	0	40
meduc	1,802	13.7	2.1	3	17
mage	1,832	29.6	4.8	16	44
male	1,832	0.5	0.5	0	1
fstmon	1,832	0.3	0.5	0	1