



## • مقدمه

متغیرهای مورد بررسی در بحث عوامل موثر در تصادفات در این پروژه شامل شرایط روشنایی، شرایط آب و هوایی، وضعیت جسمانی راننده (مصرف الکل و مواد مخدر) و سرعت می باشد. فرضیه های مورد بررسی که به شرح زیر است:

1. مصرف مواد مخدر و الکل خطر تصادفات فوتی را افزایش می دهد

2. تاریکی هوا منجر به افزایش احتمال تصادفات فوتی می شود

3. شرایط آب و هوایی نامطلوب (بارانی و برفی) خطر تصادفات فوتی را افزایش می دهد.

4. افزایش سرعت مسیر سبب افزایش شدت تصادفات فوتی می شود

## • روش تحقیق

داده های مورد استفاده از پایگاه داده FARS است. داده های تصادفات سال 2019 ایالت های کارولینای شمالی و اوهایو مورد بررسی قرار گرفته است که به منظور مدلسازی، داده هایی با توضیح دهندگی بالاتر، انتخاب شدند. در مجموع از کل متغیرها، متغیرهای صفر و یکی انتخاب شدند که بیانگر مشخصاتی از جمله سن، جنسیت، نوع تصادفات، مصرف الکل و مواد مخدر، شرایط آب و هوایی و سطح روشنایی می باشند. این داده ها به صورت دستی وارد نرم افزار استتا شدند و از آن ها خروجی لازم گرفته شد. (فایل دیتا به پیوست ارسال می گردد)

متغیرهای وابسته و مستقل به شرح زیر است:

هدف این تحقیق بررسی اثر متغیرهای مختلف بر وقوع تصادفات فوتی می باشد. به این منظور متغیرهای مختلف در مدل بارگزاری شدند و با ساخت مدل های متعدد در نهایت 5 متغیر مستقل به عنوان متغیرهای اثر گذار در وقوع تصادفات فوتی مشخص شدند و برای ساخت مدل نهایی انتخاب شدند. متغیرهای استفاده شده در مدل به شرح زیر می باشند:

1. (deaths) : این متغیر به عنوان متغیر وابسته در مدل استفاده شد و بیانگر تعداد تصادفات فوتی است. مقدار 1 این متغیر به معنی فوتی بودن تصادف می باشد
2. (drinking) : این متغیر به عنوان یکی از متغیرهای مستقل در مدل انتخاب شد و بیانگر مصرف الکل در وقوع تصادف می باشد. مقدار 1 این متغیر به معنی مثبت بودن تست الکل در زمان تصادف می باشد.
3. (drugs) : این متغیر به عنوان یکی از متغیرهای مستقل در مدل انتخاب شد و بیانگر وضعیت مصرف مواد مخدر می باشد. مقدار 1 این متغیر به معنی مثبت بودن تست مخدر می باشد.
4. (dark) : این متغیر به عنوان یکی از متغیرهای مستقل در مدل انتخاب شد و بیانگر زمان وقوع تصادف (تاریک و روشن بودن مسیر) می باشد. مقدار 1 این متغیر به معنی وقوع تصادف در شب می باشد.
5. (rain) : این متغیر به عنوان یکی از متغیرهای مستقل در مدل انتخاب شد و بیانگر وضعیت آب و هوایی مسیر می باشد. مقدار 1 این متغیر به معنی شرایط بارانی مسیر است.
6. (vspd-lim) : این متغیر به عنوان یکی از متغیرهای مستقل در مدل انتخاب شد و بیانگر وضعیت سرعت مسیر می باشد.

متغیر	نماد	واحد
متغیر مستقل	drinking	0 و 1
	drugs	0 و 1
	dark	0 و 1
	rain	0 و 1
	Vspd-lim	mph
متغیر وابسته	deaths	0-3

Variable	Obs	Mean	Std. dev.	Min	Max
deaths	1,478	.6123139	.5634163	0	3

. sum drinking drugs dark rain Vspd\_lim

Variable	Obs	Mean	Std. dev.	Min	Max
drinking	1,478	.2124493	.4091795	0	1
drugs	1,478	.2131258	.4096544	0	1
dark	1,478	.3924222	.4884552	0	1
rain	1,478	.244249	.4297865	0	1
Vspd_lim	1,453	49.73503	11.46239	0	70

مدلسازی

از مدل لجستیک برای مدلسازی استفاده شد است

گرسین لجستیک زمانی استفاده می شود که متغیر وابسته به صورت گروه بندی شده باشد، متغیر وابسته می تواند ذاتا طبقه بندی شده باشد، بعنوان مثال:

آیا تصادف رخ داده است (1) یا خیر (0)

هدف از طراحی مدل لجستیک این است که با استفاده از یک سری متغیر مستقل پیش بینی انجام دهیم، بطوریکه اگر یک فرد جدید از جامعه ای که مدل سازی را بر اساس اطلاعات نمونه ای این جامعه انجام داده ایم را انتخاب کنیم مدل پیش بینی کند که فرد در کدام طبقه قرار می گیرد.

انواع رگرسین لجستیک

1. (Binary logistic regression)
2. (Multinomial logistic regression)
3. (Ordinal logistic regression)

ما یک مدل لجستیک باینری را در تجزیه و تحلیل خود برای تعیین احتمال وقوع تصادف با توجه به ویژگی های مختلف اعمال می کنیم. در این مدل رگرسین متغیر وابسته به صورت دو وجهی می باشد لذا، تنها دو گروه خواهیم داشت.

• نتایج

جدول 1:

<b>. logit deaths drinking drugs dark rain Vspd_lim</b>						
Iteration 0: log likelihood = <b>-990.27049</b>						
Iteration 1: log likelihood = <b>-929.13599</b>						
Iteration 2: log likelihood = <b>-928.75045</b>						
Iteration 3: log likelihood = <b>-928.75035</b>						
Iteration 4: log likelihood = <b>-928.75035</b>						
Logistic regression						
Number of obs = <b>1,453</b>						
LR chi2(5) = <b>123.04</b>						
Prob > chi2 = <b>0.0000</b>						
Pseudo R2 = <b>0.0621</b>						
Log likelihood = <b>-928.75035</b>						
deaths	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
drinking	<b>.3699739</b>	<b>.1389756</b>	<b>2.66</b>	<b>0.008</b>	<b>.0975867</b>	<b>.6423611</b>
drugs	<b>1.268625</b>	<b>.1549441</b>	<b>8.19</b>	<b>0.000</b>	<b>.9649398</b>	<b>1.572309</b>
dark	<b>.2716916</b>	<b>.1138368</b>	<b>2.39</b>	<b>0.017</b>	<b>.0485756</b>	<b>.4948076</b>
rain	<b>.6192846</b>	<b>.1336965</b>	<b>4.63</b>	<b>0.000</b>	<b>.3572443</b>	<b>.8813249</b>
Vspd_lim	<b>.0115088</b>	<b>.0048435</b>	<b>2.38</b>	<b>0.017</b>	<b>.0020157</b>	<b>.021002</b>
_cons	<b>-.8246056</b>	<b>.2581962</b>	<b>-3.19</b>	<b>0.001</b>	<b>-1.330661</b>	<b>-.3185504</b>

در جدول 1 مثبت بودن ضرایب پارامترهای به دست آمده نشان دهنده‌ی احتمال بیشتر تعداد تصادفات فوتی و منفی بودن آنها نشان دهنده احتمال کمتر می‌باشد.

در این مدل با توجه به مثبت بودن تمام ضرایب تخمین زده شده، زمانی که فرد الکل استفاده کرده باشد و همچنین در صورت استفاده از مواد مخدر، تاریک بودن هوا، بارش باران و داشتن سرعت غیرمجاز تعداد فوتی‌ها افزایش می‌یابد.

## جدول 2:

```
Iteration 0:  log likelihood = -990.27049
Iteration 1:  log likelihood = -929.13599
Iteration 2:  log likelihood = -928.75045
Iteration 3:  log likelihood = -928.75035
Iteration 4:  log likelihood = -928.75035
```

```
Logistic regression                                Number of obs = 1,453
                                                    LR chi2(5)      = 123.04
                                                    Prob > chi2     = 0.0000
Log likelihood = -928.75035                        Pseudo R2      = 0.0621
```

deaths	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
drinking	1.447697	.2011946	2.66	0.008	1.102507	1.900964
drugs	3.555958	.5509746	8.19	0.000	2.62463	4.817761
dark	1.312182	.1493746	2.39	0.017	1.049775	1.640183
rain	1.857599	.2483544	4.63	0.000	1.429385	2.414096
Vspd_lim	1.011575	.0048996	2.38	0.017	1.002018	1.021224
_cons	.4384079	.1131952	-3.19	0.001	.2643026	.7272024

Note: **\_cons** estimates baseline odds.

در جدول 2 نسبت شانس (Odds ratio) همهی متغیرها مثبت و بین صفر و مثبت بی‌نهایت می‌باشد. و هر چه Odds ratio بزرگتر باشد به این معنا است که متغیر اثر بزرگتری دارد. در این مدل با توجه به نتایج نشان داده شده، بزرگترین اثر مربوط به استفاده یا عدم استفاده از مواد مخدر (drugs) و کوچکترین اثر مربوط به داشتن یا نداشتن سرعت غیرمجاز (Vspd-lim) می‌باشد.

شانس یا احتمال مست بودن نسبت به حالتی که فرد الکل مصرف نکرده است، باعث افزایش تعداد فوتی‌ها به میزان 44.76٪ می‌شود.

شانس یا احتمال مصرف مواد مخدر نسبت به حالتی که فرد استفاده نکرده باشد، باعث افزایش تعداد فوتی‌ها به میزان 255.594٪ می‌شود.

شانس یا احتمال تاریک بودن نسبت به حالتی که روز و یا روشن است، باعث افزایش تعداد فوتی‌ها به میزان 31.21٪ می‌شود.

شانس یا احتمال بارش باران نسبت به حالتی که وجود ندارد، باعث افزایش تعداد فوتی‌ها به میزان 85.75٪ می‌شود.

شانس یا احتمال عدم رعایت سرعت مجاز نسبت به حالتی که رعایت می‌شود، باعث افزایش تعداد فوتی‌ها به میزان 1.15٪ می‌شود.

value مدل عددی بسیار کوچک می‌باشد، و فرضیه اینکه تمام ضرایب برابر صفر هستند رد می‌شود و به عبارتی دیگر مدل ما به صورت کلی معنادار است، بنابراین این مدل با سطح اطمینان خیلی بالایی معنادار است.

همچنین P-value ضرایب به دست آمده نیز کوچک و کمتر از 0.05 می‌باشد و در سطح اطمینان 95٪ تمام متغیرهای موجود اثر معناداری روی تعداد تصادفات فوتی دارند.

همچنین بازه اطمینان برای ضرایب تخمین زده شده شامل صفر نیست که نشان دهنده‌ی معناداری متغیرها می‌باشد.

### جدول 3:

Akaike's information criterion and Bayesian information criterion						
Model	N	ll(null)	ll(model)	df	AIC	BIC
.	1,453	-990.2705	-928.7503	6	1869.501	1901.189

Note: BIC uses N = number of observations. See [\[R\] BIC note](#).

پارامتر BIC و AIC مدل محاسبه گردید که برای مقایسه با مدل‌های دیگر مورد استفاده قرار می‌گیرد. هرچه این دو پارامتر نسبت به مدل‌های دیگر کوچکتر باشد به معنای بهتر بودن مدل است.

#### • نتیجه گیری

نتایج مدل برازش شده نشان داد، زمانی که فرد از الکل استفاده کرده باشد (عدم تعادل و تمرکز و هوشیاری کامل) و همچنین در صورت استفاده از مواد مخدر (عدم تعادل و تمرکز و هوشیاری کامل)، تاریک بودن هوا (کاهش دید)، بارش باران (کاهش اصطکاک) و داشتن سرعت غیرمجاز (افزایش سرعت) تعداد فوتی‌ها افزایش می‌یابد.