

### ۵-۳-۱- زمان بقا:

زمان بقا یا فاصله زمانی ورود به مطالعه تا وقوع حادثه، یکی از متغیرهای مهم در تجزیه و تحلیل داده‌های بقا است. زمان بقا فرد  $i$  ام مطالعه یک متغیر تصادفی است که با  $T_i$  نشان داده می‌شود و فاصله زمانی نقطه ورود فرد  $i$  ام نمونه به مطالعه تا زمان وقوع واقعه مورد نظر مثل مرگ و ... را شامل می‌شود. این متغیر برای هر فرد متفاوت است.

### ۶-۳-۱- داده‌های سانسور شده:

ویژگی مهمی که همواره تجزیه و تحلیل داده‌های بقا را با مشکل مواجه می‌سازد، مقادیر سانسور شده است و آن به مواردی اطلاق می‌شود که در طول مدت پیگیری از دسته رفته و سانسور شده است و در حقیقت ثبت وضعیت نهایی آن برحسب وقوع واقعه مورد علاقه نیست.

بعضی اوقات قبل از اتمام دوره پیگیری و مراقبت، حادثه مورد علاقه برای برخی از افراد نمونه اتفاق نمی‌افتد و بدین مفهوم است که قبل از اینکه حادثه برای افراد نمونه اتفاق افتد دوره مشاهده سپری شده و به پایان می‌رسد. لذا مشخص نیست که پیشامد مورد نظر برای این دسته از افراد نمونه چه زمانی اتفاق می‌افتد. یا اینکه آیا اصلاً اتفاق خواهد افتاد؟ مشخص نیست که پیشامد مورد نظر برای این دسته از افراد نمونه چه زمانی

اتفاق می افتد. فقط یک چیز مشخص است و آن این که حادثه مورد نظر تا پایان مطالعه اتفاق نیفتاده است. از طرف دیگر برخی از افراد تحت بررسی ممکن است قبل از سپری شدن دوره مشاهده، حاضر به همکاری بیشتر با محقق نشوند و یا اینکه به دلایل دیگری غیر از بیماری مورد بررسی، مثل مهاجرت و یا حتی مرگ ناشی از علت دیگری، از مطالعه خارج شوند و دیگر در دسترس نباشند. به همه این موارد لفظ مقادیر سانسور شده اطلاق می گردد.

متغیر جدید  $V$  را که می تواند فقط مقادیر صفر و یک اختیار کند چنین تعریف می کنیم که اگر واقعه مورد علاقه (مثلاً مرگ) در طول دوره مطالعه اتفاق افتاد و نتیجه نهائی فرد نمونه، مشاهده گردید (یعنی  $T_i \leq c$ ) آنگاه مقدار  $V=1$  است. در غیر این صورت چنانچه در طی آن مدت، واقعه مورد نظر اتفاق نیفتاد آنگاه مقدار  $V=0$  در نظر گرفته شود فرد نمونه از آن جهت بعنوان سانسور شده در نظر گرفته می شود که نتیجه نهائی اش معلوم و مشخص نیست.

مقادیر سانسور شده در داده های بقا بسیار رایج بوده و روشهای آماری متداول چون آزمون  $t$  نمی تواند پاسخگوی آنالیز اطلاعات سانسور شده باشد از آن جا که اطلاعات سانسور شده فی نفسه ارزشمند بوده و حداقل حاوی این نکته است که تا زمان سانسور شدن فرد نمونه حادثه مورد نظر برایش اتفاق نیفتاده است بنابر این نه می توان سانسور شدن آنها را دور ریخت و نه می توان زمان سانسور شدن را با زمان وقوع حوادث و

زمان بقا یکسان پنداشت. لذا تجزیه و تحلیل آنها روشهای آماری خاصی را می‌طلبد. سانسور در اشکال مختلف و به دلایل گوناگونی می‌تواند اتفاق افتد. یکی از اشکال آن وقوع برحسب زمان بقا (متغیر  $T_i$ ) می‌باشد و می‌توان بر این مبنا، یک افتراق اصولی به شرح زیر برای آن قائل شد.

### ۱- سانسور شده از راست:

اگر از دست دادن فرد نمونه پس از ورود وی در مطالعه اتفاق افتد آن مشاهده را سانسور شده از راست می‌گویند. این نام از آنجهت است که زمان از دست دادن فرد تحت مراقبت در روی محور، در طرف راست نقطه ورود به مطالعه می‌باشد، یعنی  $T_i \geq t_0$  است.

### ۲- سانسور شده از چپ:

اگر زمان بقای واقعی فرد نمونه قبل از زمان شروع مشاهده و مطالعه باشد یعنی حادثه مورد علاقه قبل از ورود فرد در مطالعه اتفاق افتاده و او را از دست داده باشیم آن مشاهده را سانسور شده از چپ نامند. این نام بدان جهت است که زمان از دست دادن فرد تحت مراقبت، در طرف چپ نقطه ورود به مطالعه می‌باشد، یعنی  $T_i < t_0$  است. برای مثال در مطالعه‌ای به بررسی سن شروع قاعدگی در زنان می‌پردازند و بدین منظور افراد مونث از سن ۱۲ سالگی وارد مطالعه شده، برای مدت ۴ سال تحت مراقبت و

پیگیری قرار می‌گیرند. حال بعضی از این دختران نمونه، ممکن است قبل از ورود به مطالعه یعنی قبل از سن ۱۲ سالگی قاعده شده باشند که در این حالت سن شروع قاعدگی آنان سانسور شده از چپ است. (مگر آنکه بتوان اطلاعاتی برای زمان شروع قاعدگی این دختران به دست آورد.)

### ۳- سانسور شده فاصله‌ای:

در حقیقت ترکیبی از هر دو نوع سانسور چپ و راست با هم است، و زمانیست که حادثه مورد علاقه در یک فاصله زمانی بین زمان  $b, a$  بوقوع پیوندد یعنی  $a < T_i < b$  باشد آنرا سانسور شده فاصله‌ای می‌نامند.

### ۴- سانسور شده تصادفی:

به مواردی اطلاق می‌شود که توقف مشاهدات در اختیار و کنترل محقق نباشد. دلایل متعددی برای اینکار وجود دارد از آن جمله: افراد مورد بررسی به مناطق دیگری مهاجرت کرده باشند که تماس با آنها ممکن نباشد، یا از ادامه همکاری و شرکت در مطالعه منصرف شده باشند و یا حتی به علل دیگری مثلاً تصادف و ... که مدنظر مطالعه نیست فوت کرده باشند. حالت دیگر سانسور شده تصادفی وقتی است که زمان توقف مطالعه برای همه افراد نمونه، یکسان و از قبل مشخص شده باشد، ولی زمان ورود آنها به مطالعه متفاوت و تصادفی باشد. در این نوع بررسی‌ها بجای تعیین یک دوره زمانی

مشخص، تاریخ اتمام بررسی را از قبل تعیین می کنند (مثلاً ۱۳۸۷/۱۲/۲۹). بنابراین همه کسانی که در آن تاریخ زنده هستند سانسور شده تصادفی می باشند چون ورودشان به مطالعه، در اختیار و کنترل محقق نبوده است. مثلاً مطالعه ای به بررسی طلاق در بین مزدوجین می پردازد و آنها را برای مدت ۱۰ سال تحت مراقبت و پیگیری قرار می دهد. زمان ورود به مطالعه، همان زمان ازدواج آنهاست که نقطه ورود یا مبدأ زمان نامیده می شود. بدیهی است که پس از ۱۰ سال هنوز، عده ای از نمونه ها با هم زندگی می کنند که سانسور نوع I می باشند و در ذیل به بیان آن خواهیم پرداخت. در این فاصله ممکن است بعضی از زوجین (چه زن، چه مرد) قبل از اتمام ۱۰ سال پیگیری بمیرند که دیگر طلاق در مورد آنها مفهومی ندارد. بعضی دیگر از خانواده ها ممکن است به جای دیگری مهاجرت نمایند که پیگیری آنها مقدور نباشد. همچنین امکان دارد برخی از خانواده ها پس از مدتی از ادامه همکاری با مطالعه منصرف شوند و حاضر به شرکت در مطالعه نباشند. این نوع سانسورها، سانسور تصادفی نام دارد چون ورود به مطالعه مزدوجین یکسان نمی باشد.

از طرفی می توان سانسور از راست را به دو نوع اول (I) و دوم (II) تقسیم بندی نمود.

• **سانسور نوع اول:** سانسور نوع اول ساده ترین و متداول ترین نوع سانسور است و زمانی اتفاق می افتد که برای همه افراد نمونه، فاصله زمانی مراقبت، مشاهده و پیگیری یکسانی در نظر گرفته شود. در اینجا نوع اول بدین مفهوم است که زمان سانسور

افراد، تحت کنترل محقق بوده و از پیش ثابت فرض شده است. چون در این دوره پیگیری ثابت و از پیش تعیین شده است، واقعه مورد علاقه برای آنها اتفاق نیفتاده، و پس از آن نیز قابل مشاهده و ثبت نمی‌باشد، لذا بعنوان داده سانسور در نظر گرفته می‌شوند.

• **سانسور نوع دوم:** این نوع سانسور وقتی اتفاق می‌افتد که پس از مشاهده تعداد مشخصی از وقایع مورد علاقه، مطالعه متوقف گردد و برای آن، زمان معینی از پیش در نظر گرفته نشده باشد. مثلاً محقق تصمیم به اتمام مطالعه و توقف پیگیری افراد نمونه را پس از مشاهده ۵۰ مورد حادثه از ۸۰ مورد، در نظر داشته باشد. این نوع سانسور زیاد متداول نیست ولی در بسیاری موارد از نظر اقتصادی مقرون به صرفه است.

## ۵- سانسور اخباری:

مطالعه را با  $n$  ۵۰ نفر که تحت مراقبت و پیگیری هستند در نظر بگیرید که تعداد  $m$  ۵ نفر آنها در زمان  $t$  سانسور شده باشند. اگر تعداد  $r$  ۲۰ نفر از این ۵۰ نفری که تحت مراقبت و پیگیری هستند از نظر متغیرهای توجیهی دارای مقادیر یکسانی با آن ۵ نفر سانسور شده باشند سانسور آن ۵ نفر را سانسور اخباری نامند. یعنی آن ۵ نفر سانسور را زیر مجموعه‌ای از این ۲۰ نفر می‌توان در نظر گرفت. سانسور اخباری می‌تواند منجر به اریبی شدید در برآورد پارامترها گردد. اما تشخیص جهت و اندازه اریبی دشوار است. هیچگونه آزمون آماری نیز جهت تشخیص سانسور اخباری از سانسور غیر