



التیام

شاپا الکترونیکی: ۲۷۸۳۳۲۹۱

[eltiam.ivsa@yahoo.com](mailto:eltiam.ivsa@yahoo.com)

<http://eltiamjournal.ir/>

## مروری بر روش‌های تعیین تعداد نمونه، نمونه‌گیری آماری و گردآوری داده‌ها در پژوهش‌های دامپزشکی

داریوش سعادت<sup>۱\*</sup>، سمیرا سعادت‌جو<sup>۲</sup>، علی انوش<sup>۳</sup>

۱- دانشیار گروه بهداشت مواد غذایی، دانشکده دامپزشکی، دانشگاه زابل، زابل، ایران

۲- دانش‌آموخته دکتری عمومی دامپزشکی، دانشکده دامپزشکی تهران و کلینیسین دامپزشکی در زابل، زابل، ایران

۳- دانشجوی دکتری ژنتیک و اصلاح دام و طیور، دانشکده کشاورزی، دانشگاه زابل، زابل، ایران

\* [saadatdariush@uoz.ac.ir](mailto:saadatdariush@uoz.ac.ir)

تاریخ دریافت: ۱۴۰۲/۹/۴ تاریخ پذیرش: ۱۴۰۲/۹/۱۳



<https://doi.org/10.61186/eltiam.10.2.3>



کپی‌رایت © مجله التیام: دسترسی آزاد، کپی‌برداری، توزیع و نشر برای استفاده کامل با ذکر منبع آزاد است، © نویسندگان. ناشر: انجمن جراحی دامپزشکی ایران.

### چکیده

در شروع هر مطالعه‌ای، معمولاً این سؤال‌ها به ذهن پژوهشگر می‌رسد؛ چه تعداد نمونه برای انجام این مطالعه باید اخذ شود؟ این نمونه‌ها با چه روشی از بین اعضای جامعه هدف انتخاب می‌شوند؟ داده‌های مورد نیاز برای انجام این پژوهش با چه روشی باید جمع‌آوری شوند؟ در این مقاله تلاش می‌شود به این سؤال‌ها پاسخ داده شود.

برای انتخاب تعداد نمونه مناسب باید به روش آماری که قرار است داده‌ها با آن تجزیه و تحلیل شوند، توجه شود. در بررسی‌ها پژوهشگر قصد دارد با توجه به میانگین یا شیوعی که در نمونه آماری به دست آمده، میانگین یا شیوع مورد نظر را در جامعه با حدود اطمینان ۹۵ درصد برآورد نماید. دو فرمول برای تعیین حجم نمونه در بررسی‌ها وجود دارد که به آن‌ها اشاره می‌شود. در مطالعه‌ها هدف مقایسه میانگین یا شیوع بین گروه‌های مختلف است. در هر مطالعه با توجه به این که از کدام آزمون آماری برای آنالیز داده‌ها استفاده شود، فرمول خاصی برای تعیین تعداد نمونه وجود دارد که به برخی از آن‌ها در این مقاله اشاره می‌شود.

برای انتخاب روش نمونه‌گیری از لحاظ آماری باید به نوع پژوهش توجه شود، در بررسی‌ها و مطالعه‌ها مشاهده‌ای مقطعی نمونه آماری باید به صورت تصادفی از بین اعضای جامعه انتخاب شود. در غیر این صورت نمونه معرف جامعه هدف نخواهد بود. اما در مطالعه‌های مداخله‌ای، بحث روش نمونه‌گیری آماری مطرح نیست، بلکه ضوابطی برای ورود به مطالعه وجود دارد و هر فرد یا حیوانی (در

مطالعه‌های دامپزشکی) که این ضوابط را داشته باشد؛ می‌تواند وارد مطالعه شود. پس از ورود اعضاء به مطالعه مداخله‌ای، باید این اعضا به صورت تصادفی به تیمارهای مختلف تخصیص داده شوند.

داده‌های مورد نیاز برای انجام پژوهش را می‌توان از طریق معاینه، مشاهده، انجام آزمایش‌ها یا مصاحبه جمع‌آوری کرد. برای انجام مصاحبه باید پرسشنامه‌ای طراحی شود. در طراحی پرسشنامه باید به پایایی و روایی سؤال‌ها توجه شود.

**واژه‌های کلیدی:** تعیین تعداد نمونه، روش نمونه‌گیری آماری، روش جمع‌آوری داده‌ها

### تعیین تعداد نمونه (Sample Size) مناسب

گوسفندان منطقه سیستان با روش PCR انجام شد، بر اساس پژوهش‌های قبلی شیوع ملکولی بروسلوز در منطقه سیستان ۱۱ درصد برآورد شد. با پذیرش این که شیوع به دست آمده در نمونه با شیوع واقعی در جامعه حداکثر ۵ درصد اختلاف داشته باشد ( $d=0.05$ ) و در نظر گرفتن سطح اطمینان ۹۵ درصد تعداد نمونه مورد نیاز برای این پژوهش ۱۵۰ رأس گوسفند برآورد شد (۲).

$$n = \frac{z_{1-\frac{\alpha}{2}}^2 \times p(1-p)}{d^2} = \frac{1.96^2 \times 11\% (1-11\%)}{5\%^2} = 150$$

در بررسی‌هایی که هدف پژوهشگر برآورد میانگین یک متغیر کمی (عددی) در جامعه است، از فرمول زیر برای تعیین تعداد نمونه استفاده می‌شود (۳).

$$n = \frac{z_{1-\frac{\alpha}{2}}^2 \times \sigma^2}{d^2}$$

در فرمول فوق  $n$  تعداد نمونه مورد نیاز است.  $Z_{1-\frac{\alpha}{2}}$  مقدار توزیع نرمال استاندارد ( $Z$ ) با سطح اطمینان  $1 - \alpha$  است.  $\sigma$  انحراف معیار صفت در جامعه است، که بر اساس مطالعه‌های قبلی برآورد می‌شود و  $d$  دقت است، یعنی حداکثر چه مقدار خطا در برآورد میانگین قابل قبول است.

برخلاف بررسی‌ها، در مطالعه‌ها به مقایسه گروه‌های مختلف حیوانات پرداخته می‌شود (۴). یافته‌های حاصل از مطالعه با آزمون آماری تجزیه و تحلیل می‌شود. آزمون آماری با توجه به یافته‌ها به دست آمده از نمونه قضاوت می‌کند که آیا تفاوتی بین پارامترهای جامعه وجود دارد یا خیر؟ اما آزمون آماری اعلام نمی‌کند که این تفاوت‌ها چقدر است. به‌عنوان مثال آزمون آماری ممکن است نشان دهد که یک دارو در کاهش میانگین فشار خون در جامعه مؤثر است، اما نمی‌تواند نشان دهد که این دارو چقدر میانگین فشار خون افراد جامعه

برای انتخاب تعداد نمونه مناسب باید به روش آماری که قرار است داده‌ها با آن تجزیه و تحلیل شوند، توجه شود. امروزه تجزیه و تحلیل آماری داده‌ها، یک بخش ضروری از پژوهش به حساب می‌آید و بدون انجام تجزیه و تحلیل آماری، امکان تعمیم یافته‌ها به دست آمده از نمونه آماری به کل جامعه هدف امکان‌پذیر نیست. در بررسی‌ها پژوهشگر قصد دارد با توجه به میانگین یا شیوعی که در نمونه آماری به دست آورده، میانگین یا شیوع مورد نظر را در جامعه با حدود اطمینان ۹۵ درصد برآورد نماید. دو فرمول برای تعیین تعداد نمونه در بررسی‌ها وجود دارد. یکی از این دو فرمول زمانی استفاده می‌شود که هدف برآورد میانگین یک صفت در جامعه باشد و دیگری برای زمانی است که هدف پژوهش برآورد شیوع در جامعه است.

در بررسی‌هایی که هدف صرفاً تعیین شیوع یک بیماری یا آلودگی در یک منطقه است، از فرمول زیر که معروف به فرمول کوکران برای تعیین حجم نمونه استفاده می‌شود (۱).

$$n = \frac{z_{1-\frac{\alpha}{2}}^2 \times p(1-p)}{d^2}$$

در فرمول فوق  $n$  تعداد نمونه مورد نیاز است.  $Z_{1-\frac{\alpha}{2}}$  مقدار توزیع نرمال استاندارد ( $Z$ ) با سطح اطمینان  $1 - \alpha$  است. از آنجا که معمولاً سطح اطمینان ۹۵ درصد در نظر گرفته می‌شود، مقدار  $Z_{1-\frac{\alpha}{2}}$  در فرمول فوق ۱.۹۶ است.  $p$  شیوع بیماری در جامعه است، که بر اساس مطالعه‌های قبلی برآورد می‌شود.  $d$  دقت است، یعنی حداکثر چه مقدار خطا در برآورد شیوع قابل قبول است. برای مثال، برای تعیین تعداد نمونه مورد نیاز در پژوهشی که به منظور بررسی شیوع بروسلوز در

اول و دوم کاهش داده شود، باید تعداد نمونه بیشتری وجود داشته باشد، اما معمولاً مقدار خطای نوع اول (آلفا) را ۰/۰۵ و مقدار خطای نوع دوم (بتا) را ۰/۱۰ یا ۰/۲۰ در نظر می‌گیرند (۵)

در این جا به سه مورد از فرمول‌های تعیین حجم نمونه در مطالعه‌ها اشاره می‌شود.

$$n = \left( Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right)^2 \frac{P_1(1 - P_1) + P_2(1 - P_2)}{(P_1 - P_2)^2}$$

### تعداد نمونه مناسب برای مقایسه شیوع بیماری بین دو گروه (۸-۵):

در فرمول بالا  $n$  تعداد (حجم) نمونه مورد نیاز در هر گروه است.  $Z$  مقدار توزیع نرمال استاندارد با اندیس مورد نظر است. آلفا ( $\alpha$ ) و بتا ( $\beta$ ) خطای نوع اول و دوم هستند. با در نظر گرفتن ۰/۰۵ خطای نوع اول، مقدار  $Z_{1-\frac{\alpha}{2}}$  برابر ۱/۹۶ می‌شود. اگر مقدار خطای نوع دوم ۰/۱۰ یا ۰/۲۰ در نظر گرفته شود، مقدار  $Z_{1-\beta}$  به ترتیب برابر ۱/۲۸ و ۰/۸۴ است.  $P_1$  و  $P_2$  به ترتیب شیوع بیماری در جامعه ۱ و ۲ را نشان می‌دهند.

به عنوان مثال در پژوهشی که به منظور برآورد شیوع توکسوپلازما در گوشت گاوهای کشتار شده در زابل و زاهدان انجام شد، فرضیه این بود که شیوع توکسوپلازما در گوشت گاوهای وارداتی از پاکستان بیشتر از شیوع این انگل در گاوهای بومی است. بر اساس پژوهش‌های گذشته برآورد شد، شیوع در گوشت گاوهای وارداتی و بومی به ترتیب ۲۸ و ۷ درصد است. بنابراین حجم نمونه مورد نیاز برای هر گروه با توجه به فرمول یاد شده برابر ۵۰ نمونه (در مجموع ۱۰۰ نمونه) برآورد شد (۹).

$$n = \left( Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right)^2 \frac{P_1(1 - P_1) + P_2(1 - P_2)}{(P_1 - P_2)^2} = (1.96 + 0.84)^2 \frac{0.28(1 - 0.28) + 0.07(1 - 0.07)}{(0.28 - 0.07)^2} = 50$$

را کاهش می‌دهد. سؤالی که به ذهن می‌رسد این است که آیا تأثیر این دارو به اندازه‌ای بزرگ هست که بتوان از آن برای کاهش فشار خون در بیماران مبتلا به پرفشاری استفاده کرد. برای پاسخ به این سؤال باید توجه داشت که آزمون آماری همانند یک ذره‌بین اختلاف‌ها را به ما نشان می‌دهد و تعداد نمونه در آزمون‌های آماری همانند قدرت ذره‌بین است (۵). همان‌طور که برای دیدن اشیای ریزتر باید از ذره‌بین قوی‌تر استفاده کرد، در مورد آزمون‌های آماری نیز برای معنی‌دار شدن اختلاف‌های کوچک‌تر باید تعداد نمونه بیشتری را انتخاب کرد. بنابراین در ابتدای هر مطالعه باید «تفاوت مهم از دیدگاه بالینی (Clinical Significance)» تعیین شود. به‌عنوان مثال در مورد داروی کاهنده فشار خون، باید دید اگر این دارو چقدر فشار خون را کاهش دهد، از لحاظ بالینی مهم و ارزشمند به حساب می‌آید. سپس تعداد نمونه در گروه شاهد و گروه درمان به اندازه‌ای قرار داده شود که اگر اثر واقعی دارو در جامعه به اندازه «تفاوت مهم از دیدگاه بالینی» باشد. این اثر معنی‌دار می‌شود (۵). باید توجه داشت که با مقایسه فشار خون بیماران در گروه شاهد و گروه درمان و بدون انجام آزمون آماری می‌توان فهمید که آیا اثر دارو در نمونه‌های مورد آزمایش به اندازه «تفاوت مهم از دیدگاه بالینی» بوده یا خیر؟ اما برای پی بردن به اثر دارو در کل جامعه انجام آزمون آماری ضروری است.

اصولاً در آزمون‌های آماری دو فرضیه متضاد  $H_0$  و  $H_1$  وجود دارد. فرضیه  $H_0$ ، فرضیه‌ای است که می‌گوید «تفاوت مهم از دیدگاه بالینی» وجود ندارد و  $H_1$  فرضیه‌ای است که می‌گوید این تفاوت وجود دارد. در انجام آزمون آماری دو نوع خطا ممکن است رخ دهد. خطای نوع اول مربوط به زمانی است که  $H_0$  درست باشد، اما آزمون آماری آن را مردود اعلام کند، خطای نوع اول را با آلفا ( $\alpha$ ) نشان می‌دهند. خطای نوع دوم زمانی اتفاق می‌افتد که  $H_0$  غلط باشد، اما آزمون آماری این فرضیه را تأیید کند. خطای نوع دوم را با بتا ( $\beta$ ) نشان می‌دهند. در مطالعه‌ها زمانی که قرار هست تعداد نمونه تعیین شود باید مشخص شود که حداکثر مقدار خطای نوع اول و دوم قابل قبول چقدر است. اگر قرار هست مقدار خطای نوع

**تعداد نمونه مناسب برای مقایسه میانگین یک صفت کمی بین دو گروه (آزمون t) (۵-۸):**

$$n = \frac{(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})^2 (\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$$

در فرمول بالا n تعداد نمونه مورد نیاز در هر گروه است. Z مقدار توزیع نرمال استاندارد با اندیس مورد نظر است. آلفا ( $\alpha$ ) و بتا ( $\beta$ ) خطای نوع اول و دوم هستند.  $\mu_2$  و  $\mu_1$  به ترتیب میانگین صفت در جامعه ۱ و ۲ هستند و  $\sigma_2$  و  $\sigma_1$  به ترتیب انحراف معیار صفت در جامعه ۱ و ۲ هستند. مقادیر میانگین و انحراف معیار در دو جامعه با توجه به مطالعه‌های گذشته برآورد می‌شوند.

**تعداد نمونه مناسب برای مقایسه میانگین یک صفت کمی بین بیش از دو گروه (۸-۵):**

در برخی از پژوهش‌ها میانگین یک صفت بین چند گروه (بیش از دو گروه) با هم مقایسه می‌شوند. آزمون آماری که در این جا (در صورت نرمال بودن داده‌ها) مورد استفاده قرار می‌گیرد، آزمون آنالیز واریانس (ANOVA) است. معمولاً

تعداد نمونه لازم برای انجام آزمون ANOVA از بسیاری از آزمون‌های دیگر کمتر است. بنابراین زمانی که میانگین‌ها مقایسه می‌شوند، انتخاب ۳ یا ۴ تیمار به جای دو تیمار می‌تواند باعث شود تمامی تعداد نمونه لازم برای انجام پژوهش کاهش یابد، برای مثال فرض کنید، پژوهشگری می‌خواهد تأثیر یک دارو بر پارامترهای بیوشیمیایی خون را بررسی کند، اگر این پژوهشگر در پژوهش خویش دو گروه (شاهد و درمان) داشته باشد برای تجزیه و تحلیل داده‌ها باید از آزمون t استفاده کند، اما اگر این پژوهش دو تیمار دارو (یکی با غلظت کم و دیگری با غلظت زیاد) و یک گروه شاهد داشته باشد، آزمون آماری مورد استفاده وی آزمون آنالیز واریانس است، که چه بسا این پژوهش در مجموع به حجم نمونه کمتری نیاز داشته باشد. فرمول تعیین تعداد نمونه برای هر گروه در آنالیز واریانس در زیر نشان داده شده است (۷):

$$n = \frac{\lambda}{\frac{1}{\sigma^2} \sum_{i=1}^K (\mu_i - \bar{\mu})^2}$$

که مقدار  $\lambda$  از جدول شماره ۱ به دست می‌آید:

جدول ۱- مقدار  $\lambda$  در تعیین تعداد نمونه برای آنالیز واریانس

تعداد گروه‌ها (K)									
۱۰	۹	۸	۷	۶	۵	۴	۳	$\alpha = 0.01$	$\beta = 0.10$
۲۶/۱۳	۲۵/۲۲	۲۴/۲۴	۲۳/۱۹	۲۲/۰۳	۲۰/۷۴	۱۹/۲۵	۱۷/۴۳	$\alpha = 0.05$	
۱۹/۸۳	۱۹/۰۹	۱۸/۲۹	۱۷/۴۲	۱۶/۴۷	۱۵/۴۱	۱۴/۱۸	۱۲/۶۶	$\alpha = 0.01$	$\beta = 0.20$
۲۱/۴۳	۲۰/۶۴	۱۹/۷۹	۱۸/۸۸	۱۷/۸۷	۱۶/۷۵	۱۵/۴۶	۱۳/۸۹	$\alpha = 0.05$	
۱۵/۶۵	۱۵/۰۳	۱۴/۳۶	۱۳/۶۳	۱۲/۸۳	۱۱/۹۴	۱۰/۹۱	۹/۶۴		

جوجه‌ها و اکسن نیوکاسل تزریق شد و اثرات سرکه سیب بر مقدار تولید آنتی‌بادی نیوکاسل در جوجه‌های واکسن زده مطالعه شد، پیش از انجام این مطالعه و به منظور تعیین تعداد نمونه مورد نیاز با توجه به مطالعه‌های پیشین میانگین تیتراژ HI برای آنتی‌بادی نیوکاسل در تیمارهای ۳، ۲ و ۱ درصد عصاره سیر و گروه شاهد به ترتیب ۴/۷، ۴/۵، ۴/۲ و ۴ و انحراف معیار تیرهای سرمی ۰/۶۵ برآورد شد. با در نظر گرفتن ۰/۰۵ خطای نوع اول و ۰/۲۰ خطای نوع دوم، تعداد

در فرمول بالا n تعداد نمونه مورد نیاز در هر گروه است.  $\sigma$  انحراف معیار داده‌ها در گروه‌ها،  $\mu_i$  میانگین در گروه آ، م، آلفا ( $\alpha$ ) و بتا ( $\beta$ ) به ترتیب خطای نوع اول و دوم، و K تعداد گروه‌ها است و مقادیر  $\lambda$  در داخل جدول شماره ۱ نشان داده شده‌اند. برای مثال در پژوهشی که درباره اثرات سرکه سیب بر عملکرد دستگاه ایمنی جوجه‌های گوشتی انجام شد (۱۰)، سرکه سیب در غلظت‌های ۱، ۲ و ۳ درصد به جیره جوجه‌ها اضافه شد و یک گروه شاهد هم در نظر گرفته شد. به همه

باشند. سپس اعضای که وارد مطالعه شده‌اند، با روش تصادفی و با کمک قرعه‌کشی به تیمارهای مختلف تخصیص داده می‌شوند (۴).

همان‌طور که گفته شد انتخاب افراد یا حیوانات در بررسی‌ها و مطالعه‌های مشاهده‌ای مقطعی باید به‌صورت تصادفی انجام شود. روش‌های نمونه‌گیری تصادفی در زیر توضیح داده شده‌اند:

### نمونه‌گیری تصادفی ساده (Simple Random Sampling)

**(Sampling):** در این روش هر عضو جامعه شانس برابری برای انتخاب شدن در نمونه دارد. داشتن فهرست تمامی اعضای جامعه به پژوهشگر کمک می‌کند تا بتواند از بین آن‌ها با کمک جدول اعداد تصادفی یا به روش قرعه‌کشی تعدادی را انتخاب کند. برای استفاده از جدول اعداد تصادفی، به هر عضو جامعه یک عدد تخصیص داده می‌شود. سپس بر اساس این جدول تعدادی شماره انتخاب می‌شوند، هر عضو (فرد یا حیوانی) که شماره‌اش با روش تصادفی انتخاب شود، جزء نمونه‌ها قرار می‌گیرد و وارد مطالعه می‌شود (۱۲). در این روش از آن‌جا که همه اعضا شانس برابری برای انتخاب شدن دارند، بنابراین نمونه اخذ شده نماینده شایسته‌ای برای جامعه است و به راحتی می‌توان یافته‌های به دست آمده از نمونه آماری را به جامعه هدف تعمیم داد. بارزترین ایراد روش نمونه‌گیری تصادفی ساده نیاز آن به فهرست کاملی از تمام اعضای جامعه است. این فهرست معمولاً برای جمعیت‌های بزرگ و جمعیت‌های دامی در دسترس نیست. در چنین مواردی، عاقلانه‌تر است که از روش‌های نمونه‌گیری دیگر استفاده شود (۱۳).

### نمونه‌گیری تصادفی سیستماتیک (Systematic Random Sampling)

**(Random Sampling):** در این روش فواصل نمونه‌گیری ثابت است. فواصل نمونه‌گیری برابر است با تعداد افراد جامعه تقسیم بر حجم نمونه. مثلاً اگر قصد انتخاب ۲۰ فرد از بین ۱۰۰۰ فرد به‌عنوان نمونه وجود داشته باشد، فواصل نمونه‌گیری برابر  $\frac{1000}{20} = 50$  خواهد بود. نخستین نمونه باید از بین شماره‌های ۱ تا ۵۰ به‌صورت تصادفی انتخاب شود،

نمونه مورد نیاز برای این پژوهش ۱۶ قطعه جوجه در هر گروه (در مجموع ۶۴ قطعه جوجه) در نظر گرفته شد (۱۰):

$$n = \frac{\lambda}{\frac{1}{\sigma^2} \sum_{i=1}^K (\mu_i - \bar{\mu})^2} = \frac{10.91}{\frac{1}{0.65^2} ((4 - 4.35)^2 + (4.2 - 4.35)^2 + (4.5 - 4.35)^2 + (4.7 - 4.35)^2)} = 16$$

در این‌جا یادآور می‌شود که با توجه هزینه‌های پژوهش معمولاً پژوهشگران به دنبال راهی هستند تا با حجم نمونه کمتر، اما با دقت کافی از نظر آماری، پژوهش خویش را انجام دهند. یکی از طرح‌هایی که می‌توان در مطالعه‌های مداخله‌ای برای کاهش تعداد نمونه مورد استفاده قرار داد، طرح متقاطع (Crossover design) است. این طرح به ویژه زمانی که تأثیر آزمایش‌ها کوتاه‌مدت باشد و پس از مدتی بتوان حیوانات گروه درمان را به عنوان شاهد و حیوانات گروه شاهد را به عنوان درمان مورد استفاده قرار داد، مفید است. برای آشنایی بیشتر خوانندگان با این طرح، مطالعه مقاله "Understanding controlled trials crossover" در این زمینه پیشنهاد می‌شود (۱۱).

### انتخاب روش نمونه‌گیری

از لحاظ آماری نمونه‌گیری به معنی انتخاب تعدادی از اعضای جامعه است که پژوهش روی آن‌ها انجام می‌شود و یافته‌های به دست آمده از آن‌ها به کل جامعه تعمیم داده می‌شود. برای انتخاب روش نمونه‌گیری مناسب باید به نوع مطالعه توجه شود. از دید اپیدمیولوژی، انواع مطالعه‌ها به دو دسته مشاهده‌ای و مداخله‌ای تقسیم می‌شوند. مطالعه‌های مشاهده‌ای خود به دو دسته مقطعی و طولی تقسیم می‌شوند. در بررسی‌ها و مطالعه‌های مشاهده‌ای مقطعی روش نمونه‌گیری باید به‌صورت تصادفی باشد. مطالعه‌های مشاهده‌ای طولی معمولاً توسط اپیدمیولوژیست‌ها انجام می‌شود و طراحی خاصی دارد، بنابراین از ذکر روش نمونه‌گیری در مطالعه‌های طولی خودداری می‌شود. در مطالعه‌های مداخله‌ای بحث نمونه‌گیری از بین افراد جامعه مطرح نیست. بلکه یک‌سری ضوابطی (Admission Criteria) برای ورود به مطالعه وجود دارد، که افراد یا حیواناتی که وارد مطالعه می‌شوند باید این ضوابط را دارا

می‌شود.

اگر صفتی که قصد پژوهش در خصوص آن وجود دارد، یک صفت کمی باشد (مانند پارامترهای خونی که با واحدهای مشخص و با عدد اندازه‌گیری می‌شوند) و هم میانگین‌ها و هم واریانس در طبقه‌های مختلف متفاوت باشند، ممکن است برای تعیین تعداد نمونه در هر طبقه و تجزیه و تحلیل داده‌های حاصل از روش تصادفی طبقه‌ای با ایرادتی مواجهه پیدا می‌کند (برای اطمینان از این‌که واریانس‌ها در طبقه‌های مختلف تفاوتی ندارند باید از آزمون F استفاده شود) (۱۳).

### نمونه‌گیری تصادفی خوشه‌ای ( Cluster Random Sampling):

این روش در جمعیت‌های بزرگ کاربرد زیادی دارد. در این روش جامعه آماری به تعدادی خوشه تقسیم می‌شوند. معمولاً در داخل هر خوشه ناهمگنی وجود دارد، اما بین خوشه‌های همگنی مشاهده می‌شود (۱۳). مزیت این روش این است که احتیاجی نیست که فهرست تمام افراد جامعه وجود داشته باشد، فقط کافی است که فهرست تمامی خوشه‌ها را داشت و مشخص باشد که در داخل هر خوشه چه تعداد نمونه وجود دارد. فرض کنید قصد انجام نمونه‌گیری تصادفی از کندوهای زنبور عسل در یک شهرستان وجود دارد. قطعاً این که فهرست تمامی کندوهای این شهرستان را داشت و به آن‌ها شماره داد، تقریباً غیرممکن است، اما اگر فهرست زنبورستان‌های این شهرستان وجود داشته باشد و بااطلاع از این‌که در هر زنبورستان حدوداً چند کندو وجود دارد، می‌توان نمونه‌گیری خوشه‌ای انجام داد. هر زنبورستان یک خوشه در نظر گرفته می‌شود. فرض کنید که در مجموع در این شهرستان ۵۰۰۰ کندوی زنبور عسل وجود داشته باشد و برای انجام این پژوهش نیاز به ۷۰ کندو از ۱۰ خوشه باشد، ابتدا فهرست زنبورستان‌ها تهیه می‌شود. فرض کنید در زنبورستان الف، ب، ج، و د به ترتیب ۱۷۰، ۲۳۰، ۲۴۰ و ۱۲۰ کندو وجود دارد. در فهرست تهیه شده فراوانی تجمعی حساب می‌شوند. فراوانی تجمعی کندوها در زنبورستان‌های الف، ب، ج، و د به ترتیب ۱۷۰، ۴۰۰، ۶۴۰ و ۷۶۰ است. برای بقیه زنبورستان‌ها (خوشه‌ها) هم فراوانی تجمعی محاسبه می‌شوند. فراوانی تجمعی آخرین زنبورستان ۵۰۰۰ است. در نظر گرفته

نمونه‌های بعدی به فواصل ۵۰ تا پس از نمونه نخست انتخاب می‌شوند، مثلاً اگر نخستین نمونه شماره ۳ باشد، نمونه بعدی نمونه شماره ۵۳ خواهد بود. بعدی ۱۰۳، بعدی ۱۵۳ و به همین ترتیب به فواصل ۵۰ تایی نمونه گرفته می‌شود. روش نمونه‌برداری سیستماتیک برای نمونه‌گیری از خط تولید مواد غذایی مناسب است. برای مثال این روش در کارخانه کنسروسازی مفید است. به‌عنوان مثال می‌توان کنسرو شماره ۳، شماره ۵۳، شماره ۱۰۳ و... از روی خط تولید به عنوان نمونه برداشته می‌شود (۱۲). مزایای روش نمونه‌گیری تصادفی سیستماتیک این است که نمونه را به‌طور یکنواخت در بین جمعیت پخش می‌کند و انجام این روش آسان‌تر از نمونه‌گیری تصادفی ساده است. محدودیت استفاده از روش نمونه‌گیری سیستماتیک در زمانی است که فواصل نمونه‌گیری با روند دوره‌ای یک متغیر در جامعه منطبق باشد. به‌عنوان مثال فرض کنید ایرادی در دستگاه بسته‌بندی کنسروها وجود داشته باشد و این ایراد با فاصله هر ۱۰ کنسرو در خط تولید تکرار شود. در این صورت نمونه‌های اخذ شده با روش سیستماتیک دیگر تصادفی محسوب نمی‌شوند و نماینده کل محصول آن کارخانه نخواهند بود (۱۳).

### نمونه‌گیری تصادفی طبقه‌ای ( Stratified Random Sampling):

زمانی که جامعه آماری هدف از چند گروه (طبقه) تشکیل شده و گروه‌ها در مقایسه با کل جمعیت همگن‌تر باشند؛ این روش نمونه‌گیری مناسب‌تر است. برای مثال با اطلاع از این‌که فون انگل‌های دامی در مناطق کوهستانی، بیابانی و جلگه‌ای متفاوت است؛ در پژوهشی که به‌منظور شناسایی فون کرم‌های شیردان در گوسفندان مناطق مختلف ایران انجام شد، کل کشور به چهار منطقه آب‌وهوایی تقسیم و نمونه‌ها به‌صورت تصادفی طبقه‌ای از منطقه‌های مختلف اخذ شدند (۱۴).

در این روش جمعیت‌های دامی ممکن است بر اساس سن، محل، نوع پرورش، گونه، نژاد، وسعت دامداری، نوع تغذیه و... طبقه‌بندی شوند. در روش طبقه‌ای پس از این‌که شیوع بیماری در هر طبقه مشخص شد، با در نظر گرفتن فراوانی نسبی هر طبقه، شیوع بیماری در کل جمعیت محاسبه

می‌توان با استفاده از یک قطعه فنس در ۵ نقطه از سالن و در هر نقطه ۱۰ پرنده‌ای که به‌طور تصادفی در آن محل قرار گرفته‌اند را در حصار قرار داد و از آن‌ها نمونه‌گیری کرد. با این کار اختیار و اراده افراد در انتخاب نمونه از بین می‌رود و نمونه‌ها تصادفی‌تر انتخاب می‌شوند.

### روش جمع‌آوری داده‌ها

روش‌های مختلفی برای جمع‌آوری داده‌های پژوهش وجود دارد. داده‌های مورد نیاز برای پژوهش را می‌توان از طریق معاینه دام‌ها، مشاهده شرایط دامداری و انجام آزمایش‌ها جمع‌آوری کرد. همچنین با مصاحبه نیز می‌توان اطلاعاتی را راجع به دام‌ها و شرایط نگهداری آن‌ها کسب نمود. در ضمن مطالعه‌هایی وجود دارد که فقط به بررسی آگاهی، نگرش و عملکرد افراد ( Knowledge Attitude and Practice Studies) می‌پردازند. چنین مطالعاتی صرفاً با مصاحبه انجام می‌شوند.

برای انجام مصاحبه باید پرسشنامه‌ای طراحی شود. در پرسشنامه ممکن است از سؤال‌های باز یا سؤال‌های بسته استفاده شود، در سؤال‌های باز مصاحبه‌شونده می‌تواند هر مطلبی که به ذهنش می‌رسد را در جواب سؤال‌ها بنویسد، اما در سؤال‌های بسته باید در بین گزینه‌های موجود یکی را انتخاب نماید، گاهی اوقات ممکن است پاسخ وی جزء گزینه‌ها نباشد. در این صورت باید نزدیک‌ترین گزینه به پاسخ را انتخاب نماید.

تجزیه و تحلیل سؤال‌های بسته راحت‌تر از سؤال‌های باز است، اما ممکن است پاسخ شامل مواردی باشد که در گزینه‌ها وجود ندارد. سؤال‌ها باید با توجه به سطح فرهنگ و سواد اشخاص مصاحبه‌شونده طرح شوند. اگر قرار است با دامداران محلی و روستائیان مصاحبه شود، سؤال‌ها باید به زبان ساده باشند، از اسامی محلی بیماری‌ها استفاده شود و از کلمات علمی استفاده نشود.

زمانی که از پرسشنامه برای سنجش نظر یا نگرش افراد استفاده می‌شود، معمولاً سؤال‌ها به‌صورت کیفی رتبه‌ای در مقیاس‌های ۵، ۷ یا ۹ رتبه‌ای مطرح می‌شوند. تجزیه و تحلیل متغیرهای کیفی رتبه‌ای با استفاده از آمار ناپارامتریک یا

می‌شود که کندوهای ۱ تا ۱۷۰ در زنبورستان الف، ۱۷۱ تا ۴۰۰ در زنبورستان ب، ۴۰۱ تا ۶۴۰ در زنبورستان ج و ۶۴۱ تا ۷۶۰ در زنبورستان د (الی آخر) قرار دارند. حالا از میان فراوانی‌های تجمعی به‌طریق تصادفی سیستماتیک، خوشه‌ها را انتخاب می‌شوند. فواصل نمونه‌گیری برابر با  $\frac{500}{10} = 50$  است. نخستین عدد را بین ۱ تا ۵۰۰ به‌طور تصادفی انتخاب می‌شوند. مثلاً ۳۳ و بعد به فواصل ۵۰۰ تا نمونه‌های بعدی انتخاب می‌شوند (۳۳، ۵۳۳، ۱۰۳۳، ۱۵۳۳، ...). در این جا کدهای ۳۳ و ۵۳۳ در زنبورستان‌های الف و ج قرار گرفته‌اند و بنابراین این دو زنبورستان انتخاب می‌شوند. هر زنبورستانی که انتخاب شد، به آن‌جا مراجعه می‌شود و از بین کندوهای آن زنبورستان ۷ کندو به‌طریق تصادفی ساده انتخاب می‌شوند. در روش نمونه‌گیری خوشه‌ای حجم نمونه مورد نیاز بیشتر از حجم نمونه در سایر روش‌های نمونه‌گیری است (۱۲).

همان‌طور که ذکر شد فقط اگر نمونه‌ای با یکی از روش‌های نمونه‌گیری تصادفی اخذ شده باشد، می‌تواند نماینده جامعه هدف باشد. با این حال در علوم زیستی به‌ویژه دامپزشکی گاهی از اوقات انتخاب نمونه تصادفی به‌راحتی امکان‌پذیر نیست. برخی از احشام به‌ویژه در دامداری‌های صنعتی شماره بدن یا پلاک گوش دارند که از این شماره‌ها می‌توان برای انتخاب تصادفی بهره برد. همچنین اگر در یک روستا قصد نمونه‌گیری تصادفی خانه‌ها وجود داشته باشد، در صورتی که کدپستی خانه‌های روستا مشخص باشد، می‌توان از بین کدهای پستی به‌طور تصادفی تعدادی را انتخاب کرد یا اگر کدپستی‌ها هم مشخص نباشد، یک مسیری برای تردد در روستا انتخاب شود و در آن مسیر با فواصل مشخص و به‌صورت تصادفی سیستماتیک اقدام به نمونه‌گیری شود، اما برای نمونه‌گیری تصادفی از سالن پرورش طیور، گله احشام، حیوانات بدون سرپرست در شهرها یا حیوانات حیات وحش که حیوانات و ... که امکان شماره‌گذاری آن‌ها وجود ندارد، قرعه‌کشی عملاً غیر ممکن است. در این‌جا باید تلاش شود با انجام ترندهایی نمونه حاصل تا حد ممکن تصادفی باشد. مثلاً اگر قرار باشد از بین پرنده‌های یک سالن مرغداری ۵۰ پرنده به‌طور تصادفی انتخاب و از آن‌ها خونگیری شوند،



به دست آمده اختلاف قابل توجهی نداشته باشند. برای اندازه‌گیری پایایی پرسشنامه از ضریب پایایی استفاده می‌شود. اگر صفر باشد به معنای عدم پایایی و اگر مثبت یک باشد، یعنی پایایی پرسشنامه کامل است. انواع روش‌های به‌دست آوردن ضریب پایایی شامل روش آلفای کرونباخ، روش بازآزمایی، روش موازی و روش کودر ریچاردسون است. معمول‌ترین و رایج‌ترین روش سنجش پایایی محاسبه ضریب آلفای کرونباخ (Cronbach's alpha coefficient) است، برای اطمینان از پلاییایی کافی پرسشنامه، می‌توان پیش از پژوهش اصلی، یک مطالعه آزمایشی (pilot study) انجام داد و ضریب کرونباخ را در مورد پاسخ‌های افراد شرکت‌کننده در مطالعه آزمایشی اندازه‌گیری کرد. نرم‌افزار SPSS قادر است آلفای کرونباخ را محاسبه کند. آلفای کرونباخ زمانی مؤثر است که سؤال‌ها به شکل درست و نادرست تدوین نشده باشند.

### تعارض منافع

بین نویسندگان تعارض در منافع گزارش نشده است.

ضریب همبستگی اسپیرمن انجام می‌شود. برای مثال در پژوهشی که به‌منظور مقایسه تأثیر تدریس درس عملی میکروبی‌شناسی به‌صورت فشرده یا طولانی بر درصد یادگیری دانشجویان رشته داروسازی انجام شد، دانشجویان به دو دسته تقسیم شدند. به یک گروه درس عملی میکروبی‌شناسی به‌صورت فشرده و به گروه دوم این درس به‌صورت طولانی‌مدت تدریس شد. پس از اتمام جلسه‌های تدریس، پرسشنامه‌ای با سؤال‌های پنج‌گزینه‌ای در مقیاس لیکرت (کاملاً موافقم، موافقم، نظری ندارم، مخالفم و کاملاً مخالفم) به دانشجویان داده شد و نظر دانشجویان راجع به روش تدریس پرسیده شد. در نهایت با استفاده از آزمون ناپارامتریک مان‌ویتنی (Mann-Whitney) پاسخ‌های دو گروه با هم مقایسه شد (۱۵).

در بررسی‌های پرسشنامه‌ای، سؤال‌ها باید ارزیابی شوند و پایایی (Validity) و روایی (Reliability) پرسشنامه مشخص شود. روایی یا اعتبار پرسشنامه یعنی این که سؤال‌های پرسشنامه تا چه حد برای جمع‌آوری داده‌های پژوهش مناسب است. پایایی یعنی این که اگر از آن پرسشنامه در چند زمان مختلف در یک جمعیت استفاده شود، یافته‌های

### منابع

1. Cochran WG. Sampling techniques: 3<sup>th</sup> edition. John Wiley & Sons; New York, 1977.
2. Narouei M-R, Saadati D, Najimi M, Gangali H, Shah Karami F. Prevalence and risk factors of Brucella infection in sheep and goats in the Sistan region by PCR method. Journal of Zoonotic Diseases. 2022. <https://doi.org/10.22034/jzd.2022.15111>
3. Mohammad K, Malekafzali H. Statistical Methods and Health Indices. 19<sup>th</sup> edition. Jafari Novin; Tehran, 2021. 360 p.
4. Michael T, Robert C, Helen B, Peter J, Diggle, Nigel F, Keith H, Louise K, Annette O, Jan S, Hannah W. Veterinary Epidemiology: 4<sup>th</sup> edition. John Wiley & Sons; 2018. <https://doi.org/10.1002/9781118280249.fmatter>
5. Haghdoost A. Do You Want to Gain a Profound Insight into Sample Size and Statistical Power. Iranian Journal of Epidemiology, 2009;5(1):57-63.
6. Charan J, Biswas T. How to calculate sample size for different study designs in medical research? Indian Journal of Psychological Medicine. 2013;35(2):121-6. <https://doi.org/10.4103/0253-7176.116232>
7. Chow S-C, Shao J, Wang H, Lokhnygina Y. Sample size calculations in clinical research: CRC press; 2017. <https://doi.org/10.1201/9780203911341>
8. Stevenson MA. Sample size estimation in veterinary epidemiologic research. Frontiers in Veterinary Science. 2021; 7:539573. <https://doi.org/10.3389/fvets.2020.539573>
9. Anvari D, Saadati D, Nabavi R, Eskandani MA. Epidemiology and molecular prevalence of toxoplasma gondii in cattle slaughtered in Zahedan and Zabol Districts, South East of



- 
- Iran. Iranian Journal of Parasitology. 2018;13(1):114. PMID: 29963093.
10. Jahantigh M, Kalantari H, Ayda Davari S, Saadati D. Effects of dietary vinegar on performance, immune response and small intestine histomorphology in 1-to 28-day broiler chickens. *Veterinary Medicine and Science*. 2021;7(3):766-72 .  
<https://doi.org/10.1002/vms3.408>.
11. Sibbald B, Roberts C. Understanding controlled trials crossover trials .*BMJ*. 1998;316(7146):1719-20.  
<https://doi.org/10.1136/bmj.316.7146.1719>.
12. Nanjundeswaraswamy T, Divakar S. Determination of sample size and sampling methods in applied research. *Proceedings on Engineering Sciences*. 2021; 3(1):25-32.  
<https://doi.org/10.24874/PES03.01.003>.
13. Sharma G. Pros and cons of different sampling techniques .*International Journal of Applied Research*. 2017;3(7):749-52.
14. Nabavi R, Eslami A, Shokrani H, Bokaie S, Shayan P, Saadati D. Study on the prevalence, intensity and seasonal dynamics of abomasal helminths in sheep from different climatic zones of Iran .*World Appl Sci J*. 2011;12(4):441-5.
15. Rashki-ghalehnoo Z, Saadati D, Tajrobe-kar O. comparison of intensive and non-intensive microbiology courses on the learning of pharmacy students. *studies in Development of Medical Education*. 2015;12(3):545-9

**Abstract in English****An overview of sampling, sample size and data collection methods in veterinary research****Dariush Saadati<sup>1\*</sup>, Samira Saadatjou<sup>2</sup>, Ali Anusha**

1 -Associate Professor, Department of Food Hygiene, Faculty of Veterinary Medicine, University of Zabol, Zabol, Iran

2-DVM Graduated from Faculty of Veterinary Medicine, university of Tehran, Veterinary clinician in Zabol, Zabol, Iran.

3-Ph.D. student in Animal genetics and breeding, Faculty of Agriculture, University of Zabol, Zabol, Iran.

\* [saadatdariush@uoz.ac.ir](mailto:saadatdariush@uoz.ac.ir)

At the beginning of any study, these questions usually come to the researcher's mind; How many samples should be taken to conduct this study? How are these samples selected from the target population? How I can collect the required data? In this article, we are going to answer these questions.

In order to choose the appropriate sample size, attention should be paid to the statistical method with which the data is to be analyzed. In the surveys, the researcher intends to estimate the desired average or prevalence in the population with 95% confidence interval. There are two formulas for determining the sample size in surveys that are referred to. In studies, the aim is to compare the average or prevalence between different groups. In each study, according to which statistical test to use for data analysis, there is a specific formula for determining the sample size, some of which are mentioned in this article.

In order to choose the sampling method from a statistical point of view, attention should be paid to the type of research, in surveys and cross-sectional observational studies, the samples should be randomly selected from among the members of the population. Otherwise, these samples will not be representative of the target population. But in interventional studies, random sampling method is not considered. Rather, there are inclusion criteria, and any person or animal (in veterinary studies) who meets these criteria can enter the study. After the members enter the intervention study, these members should be randomly assigned to different treatments.

The required data to conduct research can be collected through examination, observation, experiments or interviews. A questionnaire must be designed for the interview. In designing the questionnaire, attention should be paid to the validity and reliability of the questions.

**Keywords:** Sample size, Statistical sampling methods, Data collection methods.