

بررسی تأثیر تعداد تکرار بر اعتبار نتایج آزمایشات کشاورزی

مسعود خزاعی^{*۱}

تاریخ دریافت: ۱۳۸۹/۱۲/۱

تاریخ پذیرش: ۱۳۹۱/۲/۲۷

چکیده

یکی از اصول اساسی در هر آزمایش تکرار تیمارهای آن است به نحوی که بدون تکرار امکان استدلال آماری در مقایسه تیمارها وجود ندارد. استفاده از تعداد تکرار کافی در آزمایش تأثیر زیادی در اعتبار علمی آن دارد، ولی در اغلب موارد دقت لازم در این خصوص به عمل نمی‌آید. عموماً انتخاب تعداد تکرار براساس سوابق آزمایشاتی که در شرایط مشابه صورت گرفته است انجام می‌شود، ولی باید توجه داشت که تعداد تکرار در هر آزمایش بستگی خاص به شرایط همان آزمایش دارد. در این تحقیق تلاش شده است تا با تأکید بر آزمایشات کشاورزی، تأثیر تعداد تکرار در آزمایش بر نتایج به دست آمده مورد بحث و بررسی قرار گیرد. نتایج نشان داده است که با افزایش نوسانات خطای آزمایش، و نیز کاهش میزان اختلافی که می‌بایست بین تیمارها پیدا شود برای دستیابی به یک حداقل سطح معنی‌داری و نیز توان قابل قبول در آزمون، یک حداقل تعداد تکرار مورد نیاز خواهد بود. حداقل تعداد تکرار مورد نیاز با کنترل نوسانات خطای آزمایش کاهش می‌یابد. لذا در آزمایشاتی که در شرایط کنترل شده، نظیر اتاقک رشد، انجام می‌شود معمولاً در شرایط یکسان تعداد تکرار کمتری مورد نیاز خواهد بود. علاوه بر این در صورت یکنواختی بیشتر مواد آزمایشی نیز، تکرار کمتری مورد نیاز خواهد بود. تعداد تکرار کم موجب کاهش توان آزمون خواهد شد و فقط از طریق بالا بردن تعداد تکرار می‌توان سطح معنی‌داری و توان آزمون را هم‌زمان در سطح بالایی حفظ نمود.

واژه‌های کلیدی: تکرار، توان آزمون، سطح معنی‌داری

مقدمه

اگر چه معمولاً تعداد تکرار برای همه تیمارها یکسان است ولی گاهی ممکن است تیمارها تکرارهای نامساوی داشته باشند. تکرارهای نامساوی از تیمارها، مستلزم استفاده از طرح‌های پیچیده‌تر است و در نتیجه آنالیز و تفسیر آن نیز مشکل‌تر می‌شود. گاهی، خصوصاً زمانی که منابع محدود هستند، استفاده از این طرح‌ها ضروری است. اگر این طرح‌ها نیز به دقت طراحی شوند، ممکن است آنالیز و در نتیجه محاسبات ریاضی آن پیچیده‌تر باشد، ولی تفسیر آن واضح خواهد بود. گاهی آزمایش فقط با یک تکرار انجام می‌شود. در این حالت اثر تیمار در مقابل تخمینی از نوسان‌های زمینه‌ای مقایسه می‌شود. این مقدار از نوسان‌ها باید در ابتدای آزمایش تعریف شود و معمولاً به وسیله برخی از ترکیبات تیمار که مورد نظر محقق نیستند (به عنوان مثال یک اثر متقابل مرتبه بالا) ایجاد می‌شوند (۴). معمولاً این آزمایش‌ها، شکلی از یک آزمایش فاکتوریل چند فاکتوره هستند که در آن اثرات متقابل مرتبه بالا پولد (جمع) می‌شوند تا تخمینی از خطای آزمایش را فراهم آورند. معمولاً اثر متقابل مرتبه بالا، معنی و مفهوم کاربردی ندارند، بنابراین ترجیح داده می‌شود که اثرهای متقابل مرتبه بالا پولد شوند و به عنوان خطا مورد استفاده قرار گیرند. مزیت این

یکی از اصول اساسی در آزمایش‌ها تکرار تیمارها است. در بیان آماری، تکرار به کاربرد یک تیمار در یک یا بیشتر واحد آزمایشی اطلاق می‌شود که گاهی موجب سردرگمی می‌گردد، زیرا استدلال می‌شود که برای داشتن یک تکرار باید از هر چیز دو تا داشته باشیم و هیچ دو واحد آزمایشی حتی اگر تیمار مشابهی برای آن‌ها بکار رود، نتیجه مشابهی نخواهند داشت. داده‌های بیولوژیکی متغیر هستند، بنابراین برای این که آزمایش از نظر آماری معتبر باشد باید شامل بیش از یک تکرار باشد. به این مهم نیز باید توجه داشت که تکرار کردن یک تیمار، بیشتر از تکرار اندازه‌گیری‌ها برای یک واحد آزمایش مورد نیاز است و برای دست یافتن به تکرار، فقط می‌بایست هر تیمار در بیش از یک واحد آزمایشی بکار رود و از هر واحد آزمایشی، حداقل یک مشاهده ثبت شود.

۱- کارشناس ارشد زراعت گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه بیرجند

*- نویسنده مسئول: (Email: mkhazaei20@yahoo.com)

به‌دست آوردن تعداد تکرار کافی، بین اندازه اختلافی که باید پیدا شود، توان آماری، سطح معنی‌داری آزمون و تعداد تیماری که قرار است مورد آزمون قرار گیرد، مصالحه صورت گیرد. به‌عنوان مثال اگر تیمارها گران و یا فقط تعداد کمی حیوان برای آزمایش موجود می‌باشد و یا محدودیت زمین و فضا وجود دارد، در این خصوص توافق و مصالحه صورت خواهد گرفت.

به‌دست آوردن یک برآورد خوب از واریانس خطا نیز خیلی حائز اهمیت است. اگر واریانس خطایی که از آنالیز آزمایش به‌دست آمده است با برآوردی که از واریانس خطا به‌منظور محاسبه تعداد تکرار مورد نیاز، بکار برده شده‌است خیلی اختلاف داشته باشد، در این صورت تعداد تکرار بکار برده شده در آزمایش کافی نبوده‌است. برای یک مجموعه از شرایط داده شده (یعنی سطح معنی‌داری و توان آزمون، نوسان‌های زمینه‌ای و اندازه اختلاف تیمارها) یک حداقل تکرار مورد نیاز خواهد بود تا بتوان اختلاف بین مقادیر میانگین نمونه‌ای تیمارها را پیدا نمود. تعداد تکرار بیش از حد نیز باعث عدم‌توانایی در کنترل عوامل و نیز صرف بی‌رویه منابع خواهد شد. تعداد تکرار خیلی زیاد نیز نه عملی است و نه از نظر آماری سود بیشتری دارد. تعداد تکرار انتخاب شده باید در حدی باشد که آزمایش هم از نظر آماری و هم از نظر اقتصادی سودمند باشد. ممکن است در یک آزمایش متغیرهای زیادی اندازه‌گیری شود، ولی معمولاً برای تعیین تعداد تکرار مورد نیاز، متغیر پاسخ اصلی مورد استفاده قرار می‌گیرد. در حالتی که تعدادی از متغیرها دارای سودمندی یکسانی باشند، متغیر بی‌ثبات‌تر برای تعیین تعداد تکرار مورد نیاز بکار برده می‌شود.

مواد و روش‌ها

اطلاعات مورد نیاز

اولین چیزی که می‌بایست برای تعیین تعداد تکرار مورد نیاز در نظر گرفت اندازه اختلافی است که باید بین تیمارها پیدا شود. این اندازه به‌وسیله اطلاعات قبلی و ابزار اندازه‌گیری مشخص می‌شود. در مرحله بعد باید تخمینی از نوسان‌های خطا را به‌دست آورد. در اغلب موارد از آزمایش‌هایی که قبلاً متغیر پاسخ را تحت شرایط مشابه بررسی نموده است، استفاده می‌کنند و برآورد نوسانات خطای آن آزمایش را بکار می‌برند. اگر هیچ سابقه‌ای از آزمایش‌های قبلی وجود ندارد، یک مطالعه آزمایشی انجام می‌شود تا این اطلاعات را فراهم نماید (۳). اگر اجرای مطالعه آزمایشی امکان‌پذیر نیست و اطلاعاتی درباره نوسانات زمینه‌ای نیز موجود نمی‌باشد و آزمایش باید تحت هر شرایطی انجام شود، در این حالت باید تعداد تکراری را که از نظر مالی و عملی امکان‌پذیر است بکار برد و نوسانات زمینه‌ای را بعد از اجرای این آزمایش تخمین زد. در این صورت عطف به ماسبق می‌شود که یا

تکنیک این است که بدون نیاز به تعداد خیلی زیادی واحدهای آزمایشی، می‌توان تعداد تیمار بیشتری را بکار برد.

تکرار دو نقش اساسی در آزمایش‌ها دارد (۱۵). اول آن که در اکثر آزمایش‌ها نمی‌توان همه جمعیت را اندازه‌گیری نمود و به‌همین دلیل یک نمونه گرفته می‌شود و با استفاده از نمونه ویژگی‌های جمعیت برآورد می‌شود. با بیشتر شدن تعداد مشاهده‌ها، تخمین بهتری از ویژگی‌های جمعیت خواهیم داشت. علاوه بر این بین واحدهای آزمایشی نوسان‌های طبیعی وجود دارد و لذا دو واحد آزمایشی که تیمارهای مشابهی را دریافت می‌کنند، پاسخ یکسانی نخواهند داشت. این نوسان‌های طبیعی به عنوان خطای آزمایشی یا نوسان‌های زمینه‌ای شناخته می‌شوند و واریانس خطا یا واریانس انحراف نامیده می‌شود. برای تعیین معنی‌داری اثر تیمار، نوسانات تیمار در برابر خطای آزمایشی، یا نوسان‌های زمینه‌ای آزمون می‌شود. بدون استفاده از تکرار، اندازه‌گیری خطای آزمایشی امکان‌پذیر نیست.

تعداد تکرار مورد نیاز تحت تأثیر دو عامل اصلی است. اول اینکه در صورت وجود نوسان‌های بیشتر در داده‌ها، تعداد تکرار بیشتری مورد نیاز خواهد بود. دوم اینکه برای پیدا کردن اختلاف‌های کوچک‌تر بین تیمارها تعداد تکرار بیشتری مورد نیاز است. وقتی که آزمایشی توان کافی برای پیدا کردن اختلاف بین تیمارها را ندارد، آزمایش اعتبار علمی نخواهد داشت. اگر آزمایشی تعداد تکرار کافی برای پیدا کردن یک مقدار اختلاف مشخص را ندارد، واقع بینانه‌تر آن است که نتیجه‌گیری شود، تعداد تکرار ناکافی است، تا اینکه نتیجه بگیریم اختلاف معنی‌دار نیست. برعکس ممکن است اختلاف‌ها از نظر آماری معنی‌دار باشد ولی به حدی کوچک باشند که هیچ ارزش عملی نداشته باشد. تعداد تکرار از فاکتورهایی است که تعیین می‌کند آیا اختلاف موجود بین تیمارها، در صورتی که اختلافی وجود داشته باشد، معنی‌دار می‌شود یا خیر.

علاوه بر این سطح معنی‌داری و توان مورد نیاز آزمون نیز، تعداد تکرار مورد نیاز را تحت تأثیر قرار می‌دهد. ارتباط بین توان آزمون و تعداد تکرارهای مورد استفاده خطی نیست و دو برابر کردن تعداد تکرارها به معنی دو برابر کردن توانایی پیدا کردن اختلاف‌ها نیست. برای یک سطح معنی‌داری داده شده، هم‌زمان با افزایش تعداد تکرار، اندازه‌ای از اختلاف بین تیمارها که قرار است پیدا شود کاهش، و در نتیجه توان آزمون افزایش می‌یابد. برای یک سطح معنی‌داری، توان و تعداد تکرار خاص نیز میزان اختلافی که می‌توان بین میانگین تیمار برای یک متغیر پاسخ پیدا کرد، به میزان نوسان‌های متغیر پاسخ بستگی دارد. نوسان‌های بزرگتر، نیاز به تعداد تکرار بیشتری دارد. در عمل نوسان‌های جمعیت را نمی‌دانیم و آن را تخمین می‌زنیم. برای یک توان و سطح معنی‌داری و میزان نوسان داده‌ها، هرچه اندازه اختلافی که باید پیدا شود کمتر باشد، تعداد تکرار مورد نیاز افزایش می‌یابد ولی وابستگی خطی نیست. گاهی ممکن است به‌منظور

مشابهی دارند. مونت گومری (۱۵) و کوکران و کوکس (۱۹۷۵) تکنیک‌های متنوع و مختلفی را ارائه داده‌اند و مقدار آمارها در آن‌ها بستگی به سطح معنی‌داری مورد نیاز و توان انتخاب شده، دارد. فرمول زیر یکی از این فرمول‌هاست که از استیل و توری (۱۹۸۰) گرفته شده و می‌توان برای حالتی که بیش از دو تیمار در یک آزمایش وجود دارد، بکار برد.

$$r = 2(Z_{\alpha/2} + Z_{\beta})^2 \left(\frac{\sigma}{d}\right)^2$$

که در آن r تعداد تکرار، α احتمال خطای نوع اول (سطح معنی‌داری آزمون)، β احتمال خطای نوع دوم (یک منهای توان آزمون)، Z ناحیه در آزمایش توزیع نرمال استاندارد شده، σ انحراف معیار خطا و d اندازه‌ای از اختلاف بین میانگین تیمارها است که باید توسط آزمایش پیدا شود.

نتایج

آزمایش زراعی

در این آزمایش متغیر پاسخ اصلی عملکرد وش بود و بنابراین تأثیر تعداد تکرار با توجه به نوسانات عملکرد وش تعیین شد. از نظر عملی اختلاف تقریباً ۱۰۰ کیلوگرم وش بین تیمارها، مهم در نظر گرفته شد. انحراف معیار خطای آزمایش ۱۱۲/۵ بود ولی انحراف معیار خطا ۱۱۲/۵ و ۵۰ در نظر گرفته شده است، تا یک دامنه برای تکرارهای مورد نیاز محاسبه شود. نتایج حاصل از این بررسی در جدول ۱ آمده است (برای درک بهتر، نتایج برای دو سطح معنی‌داری ۱ و ۵ درصد ارائه شده است).

نتایج نشان داد که با توجه به واریانس خطای آزمایش با داشتن ۴ تکرار، در سطح معنی‌داری ۵ و توان ۸۰ درصد می‌توان اختلاف ۳۲۱ کیلوگرم را بین تیمارها پیدا کرد (جدول ۱). در صورتی که در همین شرایط انحراف معیار خطا به ۵۰ کاهش یابد آزمایش توانایی پیدا کردن ۱۴۱ کیلوگرم اختلاف را خواهد داشت. حال اگر با توجه به هدف آزمایش پیدا کردن اختلاف ۱۰۰ کیلوگرم بین تیمارها مورد نظر باشد در صورتی که انحراف معیار ۱۱۲/۵ و ۵۰ باشد به ترتیب ۳۲ و ۷ تکرار مورد نیاز خواهد بود. این نشان می‌دهد که برای یافتن اختلاف‌های کوچک‌تر، به تعداد تکرار بیشتر نیاز است. در صورت ثابت بودن بقیه پارامترها اگر سطح معنی‌داری ۱ درصد مورد نظر محقق باشد لازم است توان آزمون را به ۶۰ درصد کاهش داد. این بدان معنی است که در صورت ثابت بودن تعداد تکرار با کاهش سطح معنی‌داری، بتا افزایش و در نتیجه توان آزمون کاهش می‌یابد.

حالت‌های مختلف دیگری نیز در جدول ۱ نشان داده شده است که ارتباط بین تعداد تکرار، سطح معنی‌داری، توان و میزان اختلافی که می‌بایست بین تیمارها پیدا نمود را نشان می‌دهد.

تعداد تکرار کافی بکار برده نشده است و نتایج از تعداد تکرار ناکافی به دست آمده است و یا این که نه، تعداد تکرار کافی با نتایج همراه بوده است. در نتیجه این اندازه از نوسانات زمینه‌ای یک شروع و سابقه برای آزمایش‌های آینده خواهد بود.

در آخرین مرحله باید در مورد سطح معنی‌داری و توان آزمون تصمیم‌گیری نمود. غالباً سطح معنی‌داری ۵ درصد انتخاب می‌شود که بیانگر سطح ریسک مردم در اغلب حالات است. توان نیز معمولاً ۸۰ درصد گرفته می‌شود. به محض اینکه در موارد بالا تصمیم‌گیری شد، می‌توان یک آزمون آماری مناسب را از جداول آماری انتخاب نمود و برای تعیین تعداد تکرار مورد نیاز مورد استفاده قرار داد.

داده‌های آزمایشی

در این تحقیق از سه مجموعه داده، مربوط به دو آزمایش زراعی یکی در شرایط مزرعه و دیگری در شرایط کنترل شده اتاقت رشد، و یک تحقیق در باره تولید شیر در دام استفاده شده است. این آزمایشات در مزرعه و آزمایشگاه فیزیولوژی دانشکده کشاورزی بیرجند انجام شده است.

آزمایش زراعی: در این آزمایش تأثیر چهار تراکم مختلف پنبه بر عملکرد وش در یک طرح بلوک‌های کامل تصادفی با چهار تکرار مورد بررسی قرار گرفت. تعداد تکرار با توجه به آزمایشاتی که در شرایط مشابه انجام شده بود، انتخاب شد. پس از پایان آزمایش عملکرد وش اندازه‌گیری و مورد مقایسه قرار گرفت. آزمایش در شرایط کنترل شده: در این آزمایش که در شرایط اتاقت رشد، در دمای ۲۵ درجه سانتی‌گراد انجام شد، تأثیر ۵ پتانسیل اسمزی ناشی از نمک NaCl بر روی رشد گیاهچه گندم مورد بررسی قرار گرفت. با توجه به آزمایشات در شرایط مشابه، آزمایش با چهار تکرار انجام شد. پس از پایان آزمایش وزن تر و وزن خشک گیاهچه اندازه‌گیری و مورد مقایسه قرار گرفت. آزمایش دامی: در این آزمایش، از داده‌های مربوط به تولید شیر دام‌هایی که قرار بود در آزمایشی مورد آزمون قرار بگیرند استفاده شده است.

در دو آزمایش اول، تعداد تکرار با توجه به نتایج به دست آمده از آزمایشات قبلی انتخاب شد. در پایان آزمایش واریانس خطا اندازه‌گیری و تأثیر تعداد تکرار مورد استفاده در آزمایش بر نتایج مورد بررسی قرار گرفت. در آزمایش دامی با توجه به اطلاعات موجود در باره نوسانات زمینه‌ای، تعداد تکرار انتخاب شد.

محاسبه تعداد تکرار مورد نیاز

روش‌های مختلفی برای محاسبه تعداد تکرار مورد نیاز در آزمایش وجود دارد. همه این روش‌ها مبتنی بر اصول علمی مشابه بوده و نتایج

جدول ۱- تأثیر تعداد تکرار بر نتایج آزمایش زراعی

تعداد تکرار در تیمار	سطح معنی داری %	توان %	اختلاف (کیلوگرم در هکتار)	انحراف معیار
۴	۵	۸۰	۳۲۱ *	۱۱۲/۵
۴	۵	۸۰	۱۴۱ *	۵۰
۳۲ *	۵	۸۰	۱۰۰	۱۱۲/۵
۷ *	۵	۸۰	۱۰۰	۵۰
۳۲	۱	۶۰ *	۱۰۰	۱۱۲/۵
۷	۱	۶۰ *	۱۰۰	۵۰
۴۱ *	۱	۸۰	۱۰۰	۱۱۲/۵
۱۰ *	۱	۸۰	۱۰۰	۵۰
۶ *	۱	۸۰	۳۲۱	۱۱۲/۵
۶ *	۱	۸۰	۱۴۱	۵۰

(* نشان‌دهنده پارامتری است که بر اساس سایر پارامترها محاسبه شده است.)

آزمایش در شرایط اتاقک رشد

نتایج این آزمایش نشان داد که واریانس خطای آزمایش برای وزن تر و وزن خشک گیاهچه به ترتیب برابر با ۰/۰۳۶ و ۰/۰۰۲۶ است. با توجه به نتایج به دست آمده از این آزمایش می‌توان گفت که در سطح معنی داری ۵ درصد و توان ۸۰ درصد برای وزن تر و خشک گیاهچه به ترتیب آزمون توانایی پیدا کردن اختلاف ۰/۵۴ و ۰/۰۴۶ را خواهد داشت (جدول ۲). این در حالی است که میانگین وزن تر و وزن خشک گیاهچه در زمان اندازه‌گیری به ترتیب ۲/۶۴ و ۰/۲ گرم می‌باشد. البته نتایج مقایسه میانگین تیمارها نشان داد که در سطح معنی داری ۵ درصد حداقل اختلاف معنی دار، برای وزن تر و وزن خشک به ترتیب برابر با ۰/۴۶ و ۰/۰۴ است. اختلاف موجود بین نتایج آزمون مقایسه میانگین و توانایی آزمون برای پیدا کردن اختلاف معنی دار در آزمایش، به علت احتمال بروز خطای نوع اول در آزمون مقایسه میانگین است. در آزمون‌های مقایسه میانگین بجز مقایسات مستقل سطح معنی داری مقایسات برابر با سطح معنی داری به کار برده شده، نیست و با افزایش تعداد مقایسات سطح معنی داری مقایسات افزایش می‌یابد.

نتایج حاصل از این بررسی در جدول ۲ آمده است. به منظور پوشش دادن حالات بیشتر، نتایج برای دو سطح معنی داری ۱ و ۵ درصد ارائه شده است.

نتایج این بررسی نشان می‌دهد که با افزایش تکرار از ۳ به ۵ برای وزن خشک در سطح معنی داری ۵ و توان ۸۰ درصد میزان اختلافی که می‌توان بین تیمارها پیدا کرد از ۰/۰۵۷ به ۰/۰۳۹ گرم کاهش می‌یابد. این موضوع بیانگر این است که در یک شرایط

یکسان از نظر توان، سطح معنی داری، با افزایش تعداد تکرار میزان اختلافی که می‌توان بین تیمارها پیدا کرد کاهش و در نتیجه احتمال بروز خطای نوع دوم کاهش می‌یابد.

برای وزن تر گیاهچه با ۵ تکرار با کاهش سطح معنی داری آزمون از ۵ به ۱ درصد و با توان ۸۰ درصد، میزان اختلافی که می‌توان بین تیمارها پیدا کرد از ۰/۴۶ به ۰/۵۷۸ افزایش می‌یابد. در سطح معنی داری ۱ با توان ۵۳ درصد می‌توان اختلاف ۰/۴۶ گرم را بین تیمارها پیدا کرد. بنابراین در یک تعداد تکرار ثابت با افزایش سطح معنی داری توان آزمون کاهش می‌یابد. در صورتی که با توجه به هدف آزمایش، پیدا کردن ۰/۴۶ گرم اختلاف بین تیمارها مورد نظر محقق باشد در سطح معنی داری ۱ درصد و توان ۸۰ درصد استفاده از تعداد ۸ تکرار، یعنی ۴ تکرار بیشتر مورد نیاز می‌باشد.

آزمایش دامی

در این تحقیق قرار بود آزمایشی به منظور بررسی تأثیر سه رژیم غذایی بر تولید شیر در دام انجام شود. متغیر پاسخ اصلی تولید شیر در دام بود. لذا تعداد تکرار هر رژیم غذایی با توجه به نوسانات تولید شیر دامهایی که قرار بود در آزمایش وارد شوند تعیین شد.

از نظر عملی اختلاف تقریباً ۹-۷ کیلوگرم در تولید شیر روزانه بین سه رژیم غذایی، مهم در نظر گرفته شد. انحراف معیار تولید شیر دامها قبل از وارد شدن در آزمایش بین ۷-۵ کیلوگرم در روز تخمین زده شد. در این آزمایش نیز یک دامنه از اطلاعات مورد نیاز بکار برده شد تا یک دامنه از حالتها و انتخابها را پوشش دهد. نتایج حاصل از این بررسی در جدول ۳ ارائه شده است.

جدول ۲- ارتباط تعداد تکرار با توانایی پیدا کردن اختلاف در آزمایش جوانه زنی در اتاقک رشد

انحراف معیار	اختلاف (گرم)	توان	سطح معنی داری	تعداد تکرار	
۰/۰۳۶	* ۰/۶۷۳	۸۰	۵	۳	وزن تر
۰/۰۳۶	* ۰/۵۴	۸۰	۵	۴	
۰/۰۳۶	* ۰/۴۶	۸۰	۵	۵	
۰/۰۰۰۲۶	* ۰/۰۵۷	۸۰	۵	۳	وزن خشک
۰/۰۰۰۲۶	* ۰/۰۴۵۷	۸۰	۵	۴	
۰/۰۰۰۲۶	* ۰/۰۳۹۳	۸۰	۵	۵	
۰/۰۳۶	* ۰/۸۸۵	۸۰	۱	۳	وزن تر
۰/۰۳۶	* ۰/۶۸۴	۸۰	۱	۴	
۰/۰۳۶	* ۰/۵۷۸	۸۰	۱	۵	
۰/۰۰۰۲۶	* ۰/۰۷۴۹	۸۰	۱	۳	وزن خشک
۰/۰۰۰۲۶	* ۰/۰۵۷۹	۸۰	۱	۴	
۰/۰۰۰۲۶	* ۰/۰۴۸۹	۸۰	۱	۵	

(* نشان دهنده پارامتری است که بر اساس سایر پارامترها محاسبه شده است.)

جدول ۳- ارتباط تعداد تکرار ، توان و سطح معنی داری در آزمایش تولید شیر در دام

انحراف معیار	اختلاف (کیلوگرم در روز)	توان %	سطح معنی داری %	تعداد تکرار
۵	۷	۸۰	۵	۱۳ *
۷	۷	۸۰	۵	۲۳ *
۵	۹	۸۰	۵	۸ *
۷	۹	۸۰	۵	۱۵ *
۵	۷	۹۰	۱	۲۲ *
۷	۷	۹۰	۱	۴۰ *
۵	۹	۹۰	۱	۱۴ *
۷	۹	۹۰	۱	۲۵ *
۵	۷	۹۰	۵	۱۶ *
۷	۷	۹۰	۵	۳۰ *
۵	۹	۹۰	۵	۱۰ *
۷	۹	۹۰	۵	۱۹ *

(* نشان دهنده پارامتری است که بر اساس سایر پارامترها محاسبه شده است.)

خواهد بود. حال آنکه با همین سطح معنی داری و توان آزمون در صورتی که انحراف معیار ۷ باشد برای پیدا کردن ۹ کیلوگرم اختلاف در تولید شیر روزانه نیاز به ۱۵ تکرار از هر تیمار خواهد بود. همان طوری که مشاهده می شود با تغییر سطح معنی داری آزمون از ۵ به ۱ درصد و نیز توان آزمون از ۸۰ به ۹۰ درصد در شرایط ثابت از نظر انحراف معیار و میزان اختلافی که می بایست پیدا نمود تعداد تکرار بیشتری مورد نیاز است. ولی باید توجه داشت که این ارتباط خطی نیست. در آزمایشات می بایست بین این پارامترها مصالحه نمود و با توجه به امکانات و شرایط موجود انتخاب را انجام داد.

از جدول ۳ چنین به نظر می آید که برای یافتن حداقل ۷ کیلوگرم اختلاف در مقادیر میانگین تولید شیر روزانه با توان ۸۰ درصد و سطح معنی داری ۵ درصد، اگر انحراف معیار خطا ۵ کیلوگرم در روز باشد حداقل برای هر رژیم غذایی ۱۳ گاو مورد نیاز است. برای یافتن همین میزان اختلاف با توان ۹۰ درصد و سطح معنی داری ۱ درصد و انحراف معیار ۵ در هر رژیم نیاز به ۹ گاو اضافی خواهد بود.

جدول ۳ نشان می دهد که در سطح معنی داری ۵ درصد و توان ۸۰ درصد در صورتی که انحراف معیار خطای آزمایش ۵ باشد برای پیدا کردن ۹ کیلوگرم اختلاف در تولید شیر روزانه نیاز به ۸ تکرار

بحث

شایان ذکر است که گاهی ممکن است حداقل تفاوت معنی‌دار محاسبه شده در آزمایش با حداقل اختلافی که در جداول بالا محاسبه گردیده است تفاوت داشته باشد. این اختلاف ناشی از خطایی است که با آزمون‌های مقایسه میانگین همراه است و تنها در مقایسات متعامد یا مستقل سطح معنی‌داری مقایسات با سطح معنی‌داری بکار برده شده در آزمون برابر است. در صورتی که مقایسه‌ها، به مقایسه‌های مستقل محدود نشوند هم‌زمان با افزایش تعداد مقایسه‌های انجام شده بسته به نوع آزمونی که برای مقایسه میانگین تیمارها بکار برده می‌شود سطح معنی‌داری مقایسات افزایش می‌یابد. در شرایط یکسان، تعداد تیمار بکار برده شده در آزمایش نیز بر نتایج آزمایش تأثیر می‌گذارد که به دلیل جلوگیری از پیچیده شدن بحث از آن صرف‌نظر شد. فقط باید به این نکته توجه داشت که در شرایط ثابت با افزایش تعداد تیمار می‌بایست به منظور کنترل خطای نوع اول، سطح معنی‌داری مقایسات را کاهش داد.

همان‌طور که مشاهده می‌شود، بایستی در مورد تعداد تکرار، سطح معنی‌داری، توان آزمون و میزان اختلافی که می‌بایست بین تیمارها پیدا نمود، قبل از شروع آزمایش و در مرحله طرح‌ریزی تصمیم‌گیری نمود. برای این منظور می‌توان توافق‌هایی به عمل آورد و جداولی مانند جداول بالا طرح نمود. چنین جداولی اهمیت به دست آوردن برآوردهای صحیح و دقیقی از میانگین کل و نوسان‌های خطای آزمایش را نشان می‌دهد. در صورت بالاتر بودن میزان نوسان‌ها با توان، سطح معنی‌داری و اختلاف ثابت، تعداد تکرار بیشتری مورد نیاز خواهد بود. در نهایت بایستی بر این موضوع نیز تأکید نمود که تعداد تکرار در هر آزمایش بستگی خاص به شرایط همان آزمایش دارد و پس از اتمام آزمایش باید با توجه به واریانس خطای آزمایش نتایج را مورد ارزیابی قرار داد. این کار می‌تواند شروعی برای آزمایشات بعدی در شرایط مشابه باشد.

منابع

- 1- Borenstein, M., H. Rothstein, and J. Cohen. 1997. Power and Precision. Biostat. Teaneck. NJ. Software for MS-DOS systems.
- 2- Castelloe, J. 2000. Sample Size Computations and Power Analysis with the SAS System, in Proceedings of the Twenty-Fifth Annual SAS User's Group International Conference. Cary. NC. SAS Institute. Inc. Paper 265-25.
- 3- Chatfield, C. 1995. Problem solving: A Statistician Guide. second edition. Chapman and holl. London. appendix.
- 4- Cohron, W. G., and G. M. Cox. 1957. Experimental design. Second edition. John Wiley. New York. Chapters 1,2,4 and 6.
- 5- Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. Academic Press. New York. 2nd edn.
- 6- Desu, M. M., and D. Raghavarao. 1990. Sample Size Methodology. Academic Press. Boston.
- 7- Elashoff, J. 2000. Statistical Solutions. Cork. Ireland. Software for MS-DOS systems.
- 8- Freiman, J. A., T. C. Chalmers, J. H. Smith, and R. R. Kuebler. 1986. The Importance of Beta, the Type II Error, and Sample Size in the Design and Interpretation of the Randomized Controlled Trial: Survey of 71 Negative Trials, in Medical Uses of Statistics. eds. J. C. Bailar III and F. Mosteller, chap. 14: 289-304.
- 9- Hintze, J. 2000. Number Cruncher Statistical Systems. Kaysville. UT. Software for MS-DOS systems. 10
- 10- Hoenig, J. M., and D. M. Heisey. 2001. "The Abuse of Power: The Pervasive Fallacy of Power Calculations in Data Analysis," The American Statistician. 55: 19-24.
- 11- Kraemer, H. C. and S. Thiemann. 1987. How Many Subjects? Statistical Power Analysis in Research. Sage Publications. Newbury Park. CA.
- 12- Lenth, R. V. 2000. Java applets for power and sample size, <http://www.stat.uiowa.edu/~rlenth/Power/>.
- 13- Lipsey, M. W. 1990. Design Sensitivity: Statistical Power for Experimental Research. Sage Publications. Newbury Park. CA.
- 14- Mace, A. E. 1964. Sample-size determination, Reinhold. New York. Modelling Survival Data in Medical Research. by D. Collet. Chapman and Hall. 1994.
- 15- Montgomery, D. G. 1991. Design and analysis of experiments. third edition. John Wiley. New York. Chapters 3-7 and 14.
- 16- Muller, K. E. and V. A. Benignus. 1992. Increasing scientific power with statistical power, Neurotoxicology and Teratology. 14: 211-219.
- 17- O'Brien, R. G. 1998. UnifyPow.sas Version 98.08.25, Department of Biostatistics and Epidemiology. Cleveland Clinic Foundation. Cleveland. OH. Available for download from <http://www.bio.ri.ccf.org/power.html>.
- 18- Odeh, R. E. and M. Fox. 1991. Sample Size Choice: Charts for Experiments with Linear Models. Marcel Dekker. New York. second edn. Sample Size Calculation for Complex Clinical Trials with Survival Endpoints. by Joanna H. Shih. Controlled Clinical Trials 16: 395-407, 1995.

- 19- Russell, V. L. 2001. Some Practical Guidelines for Effective Sample-Size Determination Department of Statistics University of Iowa.
- 20- Schuirmann, D. 1987. A compromise test for equivalence of average bioavailability, ASA Proceedings of the Biopharmaceutical Section. 137–142.
- 21- Shuster, J. J. 1990. CRC Handbook of Sample Size Guidelines for Clinical Trials. CRC Press. Boca Raton.
- 22- Steel, R. G. D and J. H. Torrie. 1980. Principles and procedure of statistics. McGraw-Hill, New York, chapters 7-9, 15, 16 and 24.
- 23- Taylor, D. J. and K. E. Muller. 1995. “Computing Confidence Bounds for Power and Sample Size of the General Linear Univariate Model,” The American Statistician. 49: 43–47.
- 24- Thomas, L. 1997. Retrospective Power Analysis, Conservation Biology. 11: 276–280.
- 25- Thomas, L. 1998. Statistical power analysis software, [http://www. forestry.ubc.ca /conservation/power/](http://www.forestry.ubc.ca/conservation/power/).
- 26- Wright, T. 1997. A simple algorithm for tighter exact upper confidence bounds with rare attributes in finite universes, Statistics and Probability Letters. 36: 59–67.