

بررسی همبستگی‌های بین مشخصه‌های ظاهری ارقام پسته با استفاده از تحلیل عاملی اکتشافی (کلیدی)

محمدرضا دهقانی

عضو هیأت علمی دانشکده کشاورزی دانشگاه ولی عصر (عج) رفسنجان

چکیده

با وجود تنوع بالای ارقام پسته در ایران به خصوص در استان کرمان، هیچ‌گونه بررسی و پژوهشی درباره‌ی ماهیت همبستگی بین مشخصه‌ها انجام نشده است. در طرح و اجرای برنامه‌های کلاسیک به نژادی بر مبنای انتخاب ارقام برتر، اتکا به همبستگی‌های ساده بین مشخصه‌ها از توجیه علمی برخوردار نیست زیرا همبستگی‌های ساده موجود بین مشخصه‌ها دارای تفسیر روشن و واضحی نیستند. در عین حال، وجود این همبستگی‌ها می‌تواند مبین وجود روابط ویژه و عامل‌های مشترک بین مشخصه‌ها باشد. در بررسی حاضر از روش تحلیل عاملی اکتشافی برای کشف و تعیین عامل-های مشترک یا مشخصه‌های پنهان که موجب همبستگی بین مشخصه‌های قابل مشاهده می‌شوند، استفاده شد. بر این اساس، با استفاده از روش تحلیل مؤلفه‌های اصلی و دوران پرومکس تعداد ۴ مشخصه‌ی پنهان که بیشترین واریانس مشخصه‌های مورد بررسی را بیان می‌کردند، استخراج و نام‌گذاری شدند. این ۴ مشخصه روی هم $73/543$ درصد تغییرات مشخصه‌های مورد بررسی را توضیح دادند. چون شایستگی برازش مدل تحلیل عاملی با $df = 41$ تأیید نشد ($\chi^2 = 107/42$ ، $P - value = 0/0000$)، امکان بیان روابط ۴ مشخصه‌ی پنهان با مشخصه‌های آشکار در قالب یک مدل خطی به عنوان تقریبی از روابط حقیقی فراهم نشد.

مقدمه

درخت پسته اهلی *Pistacia Vera L.* متعلق به تیره سماق *Anacardiaceae* است. جنس *Pistacia* دارای ۱۱ گونه است. گیاهان این تیره به صورت درخت یا درختچه هستند. پسته به عنوان مهم‌ترین محصول باغی کشور، تولید آن دارای سابقه طولانی است. حدود ۶۵ درصد پسته جهان در ایران تولید می‌شود. شمال استان کرمان با دارا بودن بیش از ۸۰ درصد سطح زیر کشت درختان بارور، مهم‌ترین تولید کننده پسته کشور محسوب می‌شود. اقتصاد این بخش از استان کرمان به طور مستقیم یا غیر مستقیم متأثر از کمیت و کیفیت تولید این محصول است. به این ترتیب نقش تولید این محصول در صادرات غیر نفتی کشور آشکار می‌شود.

مانند سایر محصولات کشاورزی، به دلیل محدودیت جدی و بیش از پیش منابع آب و خاک زراعی، امکان افزایش تولید از طریق افزایش سطح زیر کشت بسیار اندک و فاقد توجیه اقتصادی و فنی است. بنابراین انجام بررسی‌ها و پژوهش-های علمی و فراگیر در زمینه‌ی چگونگی افزایش عملکرد در واحد سطح ضروری به نظر می‌رسد. یکی از راه‌های افزایش عملکرد در واحد سطح، طرح و اجرای برنامه‌های به نژادی و اصلاح ارقام موجود است. در طرح و اجرای برنامه‌های کلاسیک به نژادی بر مبنای انتخاب ارقام برتر، اتکا به همبستگی‌های ساده بین مشخصه‌ها از توجیه علمی و

کافی برخوردار نیست. زیرا همبستگی های ساده موجود بین مشخصه ها دارای تفسیر روشن و واضحی نیستند. در عین حال وجود این همبستگی ها می تواند مبین وجود روابط ویژه و عامل های مشترک بین مشخصه ها باشد. اسماعیل پور (۱۳۷۷) برخی از صفات کمی و کیفی ۲۸ رقم از ارقام کلکسیون مؤسسه تحقیقات پسته را مورد بررسی قرار داد. اعلمی (۱۳۷۵) گزارش کرد بین ارقام پسته رایج کشور که همگی متعلق به گونه ی *P. Vera* هستند، تنوع ژنتیکی فراوانی دیده می شود.

وارگاس و همکاران (۱۹۹۸) صفات آگرونومیکی مهم ارقام تجاری جمع آوری شده از ۱۱ کشور دنیا را مورد بررسی قرار دادند و گزارش کردند ارقام *Agina*، *Batoury* و *Mateur* زود گل و ارقام *Bianca regina* و *Pignaton* دیر گل هستند، همچنین ارقام *Mateur* و *Ouleimy* پر رشد و ارقام *Kerman* و *White ouleimy* رشد کمتری دارند. پاریت و همکاران (۱۹۹۹) صفات مهم برای انتخاب ارقام برتر را زمان گل دهی، اندازه میوه و زمان برداشت گزارش کردند. مرتضوی (۱۳۷۶) تعداد ۸۰ فنوتیپ پسته را در مناطق پسته کاری استان سمنان مورد بررسی قرار داد و آن ها را از نظر نوع مصرف به دو گروه خشکباری و تازه خوری تقسیم کرد.

کاروز و همکاران (۱۹۹۵) تنوع ژنتیکی و فنوتیپی ژرم پلاسما پسته نواحی مدیترانه ای را طی یک برنامه پژوهشی مشترک با انستیتو منابع ژنتیکی گیاهی بین المللی (*IPGRI*) بررسی نمودند. در این پژوهش بیش از ۵۰ درخت نر و ماده که از نواحی مدیترانه ای و کشورهای مختلف جمع آوری شده بودند از نظر صفات مورفولوژیک کمی و کیفی مورد بررسی قرار گرفت. آن ها از مقایسه صفات مورفولوژیک برای شناسایی ژرم پلاسما و تنوع ژنتیکی پسته استفاده کردند. الکساندر (۱۹۸۴) در انتخاب واریته های پسته ماده، عملکرد بالا و اندازه پسته بزرگ تر و درصد خندانی بالاتر را ملاک عمل قرار داد.

ماگز (۱۹۸۱) در معرفی رقم *Sirora* از مشخصه های منشأ رقم، پایه و والدین آن، عادت رشد، قدرت رشد، گل دهی، شکل برگ (نسبت طول به عرض)، موقعیت گل آذین، تاریخ رسیدگی، رنگ میوه، شکل میوه (نسبت طول به عرض)، درصد خندانی، عادت باروری و ریزش جوانه استفاده کرد.

با وجود تنوع بالای ارقام پسته در ایران به خصوص در استان کرمان، هیچ گونه بررسی و پژوهشی درباره ماهیت همبستگی های بین مشخصه ها انجام نشده است. مطالعه های انجام شده به طور عموم تا حد محاسبه ی همبستگی ها و گزارش همبستگی های معنی دار بین برخی از مشخصه ها بوده است. وجود همبستگی بین مشخصه ها می تواند ناشی از وجود عامل های مشترک یا مشخصه های پنهان باشد. در بررسی حاضر از روش تحلیل عاملی اکتشافی برای کشف و تعیین مشخصه های پنهان که موجب همبستگی های بین مشخصه های قابل مشاهده می شوند، استفاده شد. استیونس (۱۹۹۶) می گوید: تحلیل عاملی اکتشافی زمانی به کار می رود که شواهد کافی قبلی و پیش تجربی برای تشکیل فرضیه درباره ی عامل های زیربنایی مشاهدات کمی وجود نداشته باشد. در حقیقت مایل هستیم درباره ی تعیین تعداد یا ماهیت عامل هایی که همبستگی های بین مشخصه ها را توجیه می کنند، تحقیق و بررسی کنیم. بنابراین تحلیل اکتشافی بیشتر به عنوان یک روش تدوین و تولید نظریه، و نه یک روش آزمون نظریه در نظر گرفته می شود. ترستون (۱۹۳۵) پس از بررسی روش های آماری برای مطالعه اعتبار مدل های ساختاری، تحلیل عاملی اکتشافی را یک روش رسمی اکتشاف بیان کرد که فرد را قادر می سازد روابط بین متغیرهایی که هرگز در داده های اصلی یا حتی در همبستگی های بین متغیرها آشکار نیست، مشاهده کند.

استفاده از این روش در مراحل اولیه برنامه های به نژادی می تواند مفید و مؤثر باشد. در این باره انتخاب و اندازه گیری مشخصه های قابل مشاهده به عنوان نشانگرهای مشخصه های پنهان بایستی با دقت زیادی انجام شود تا اطلاعات کامل تر و بیشتری درباره ی مشخصه های پنهان فراهم شود. به این ترتیب علاوه بر مشخص شدن ماهیت همبستگی های موجود بین مشخصه ها، می توان مشخصه های قابل مشاهده را در چند مشخصه ی محدود خلاصه کرد. نتایج این بررسی به شناخت بهتر ماهیت روابط بین مشخصه ها و عوامل ایجاد همبستگی ها کمک می کند. بدیهی است این نتایج برای طرح و پیش برد برنامه های اصلاحی و به نژادی کلاسیک مفید هستند.

مواد و روش ها

برای انجام این بررسی از ۸۰ رقم پسته، شناسایی شده در استان کرمان که همه ی مشخصه های کمی و کیفی آن ها به روش های معین به عدد تبدیل و ثبت شده بود، استفاده شد. از این تعداد سه رقم به دلیل عدم قرابت زیاد مشخصه ها با سایر ارقام حذف شدند. برای ۷۷ رقم باقی مانده که شامل همه ی رقم های تجاری بودند، تعداد ۲۸ مشخصه ی کمی انتخاب گردید. این مشخصه ها عبارت بودند از: ارتفاع درخت، عرض درخت، محیط تنه، قطر تنه، طول محور گل آذین، طول شاخه رشد سال جاری، قطر وسط شاخه رشد سال جاری، طول برگ، عرض برگ، طول برگچه، عرض برگچه، مساحت برگ، طول دم برگ، قطر دم برگ، طول جوانه گل، عرض جوانه گل، وزن تر جوانه گل، طول خوشه میوه، عرض خوشه میوه، قطر دم خوشه، قطر دم میوه، تعداد انشعابات اولیه خوشه، تعداد انشعابات ثانویه روی انشعاب اولیه، وزن خوشه میوه تر، وزن محور خوشه میوه تر، وزن پسته های خندان خوشه، تعداد پسته های خوشه و اُنس خندان.

با استفاده از نرم افزار SPSS - 15 برای ۲۸ مشخصه ی گفته شده ماتریس ضرایب همبستگی محاسبه شد. با توجه به مقادیر همبستگی های به دست آمده، تعداد ۱۶ مشخصه که دارای همبستگی های بالا و معنی داری با یکدیگر بودند، انتخاب شدند. ۱۶ مشخصه انتخاب شده عبارت بودند از: ارتفاع درخت، محیط تنه، قطر تنه، قطر وسط شاخه رشد سال

جاری، طول برگ، عرض برگ، طول برگچه، عرض برگچه، مساحت برگ، طول دم برگ، طول خوشه میوه، عرض خوشه میوه، قطر دم خوشه، وزن خوشه میوه تر، وزن محور خوشه میوه تر و تعداد پسته های خوشه. وجود همبستگی های بالا بین ۱۶ مشخصه ی گفته شده می تواند ناشی از وجود عامل های مشترک یا مشخصه های پنهان باشد. در صورت امکان برای تعیین مشخصه های پنهان و بیان روابط آنها با مشخصه های مشاهده شده به عنوان نشانگر، از روش تحلیل عاملی اکتشافی استفاده شد. برای این کار ابتدا با استفاده از روش مؤلفه های اصلی تعداد ۱۶ عامل یا مشخصه ی پنهان معین شد. سپس برای رسیدن به یک ساختار ساده و قابل تفسیر با برازش مناسب از روش دوران پرومکس استفاده شد. به این ترتیب تعداد ۴ مشخصه ی پنهان که بیشترین واریانس مشخصه های مورد بررسی را بیان می کردند، استخراج شدند. با توجه به بارهای عاملی مشخصه های آشکار یا نشانگرها روی مشخصه های پنهان استخراج شده، این مشخصه ها نام گذاری شدند. برای بیان روابط مشخصه های پنهان و مشخصه های آشکار از یک مدل ساختاری عاملی استفاده شد. فرض بسیار مهم خطی بودن روابط در مدل ساختاری عاملی یک فرض متعارف و ذاتی است. برای تعیین ضرایب مسیرهای مدل و آزمون کفایت مدل از نرم افزار Lisrel - 8/5 همراه با برنامه های مربوط استفاده شد.

نتایج و بحث

جدول ۱ مقادیر همبستگی های ساده بین ۲۸ مشخصه مورد بررسی را نشان می دهد. بررسی این جدول نشان داد ۱۶ مشخصه دارای همبستگی های بالا و معنی داری هستند مقادیر همبستگی های بین این مشخصه ها در جدول ۲ خلاصه شده است. مقادیر $KMO=0/728$ و $\chi^2=493/797$ مربوط به آزمون بارتلت، کفایت داده ها را برای انجام تحلیل عاملی و هم گرائی با در نظر گرفتن حجم نمونه " $n=77$ " به نسبت، مطلوب نشان داد.

برای انجام تحلیل عاملی داده ها ابتدا با استفاده از روش مؤلفه های اصلی به تعداد مشخصه های مورد بررسی، عامل موقت استخراج شد. جدول ۳ میزان مشارکت عامل های موقت را در بیان واریانس مشخصه های مورد بررسی در مرحله -ی شروع و استخراج عامل ها نشان می دهد. کل واریانس توضیح داده شده به وسیله ی عامل های موقت و عامل های استخراج شده پس از دوران در جدول ۴ نشان داده شده است. چهار عامل اول روی هم ۷۳/۵۴۳ درصد واریانس مشخصه های مورد بررسی را توضیح می دهند. درصد کل واریانس توضیح داده شده توسط عامل ها، بعد از دوران تغییر نمی کند ولی سهم عامل های ساختار ساده نسبت به قبل از دوران متعادل شده است. بر اساس نتایج این جدول تعداد عامل ها برابر ۴ عامل انتخاب شد. شکل ۱ نمودار *Scree* را برای انتخاب تعداد عامل ها نشان می دهد. مطابق این نمودار نیز انتخاب تعداد ۴ عامل برای توضیح درصد قابل ملاحظه ای از تغییرات مشخصه ها کفایت می کند.

جدول ۵ بارهای عاملی را برای عامل های موقت نشان می دهد. مشخصه های آشکار یا نشانگرها روی عامل های موقت دارای بارهای عاملی با پراکنش زیادی هستند. این وضعیت کشف و تفسیر عامل های مشترک یا مشخصه های پنهان را مشکل می کند. برای رسیدن به یک ساختار ساده و قابل تفسیر از روش دوران پرومکس استفاده شد. جدول ۶ عامل های استخراج شده را بعد از ۶ مرتبه دوران نشان می دهد. مطابق این جدول مشخصه های آشکار یا نشانگرهای طول برگ، عرض برگ، طول برگچه، عرض برگچه و مساحت برگ دارای بارهای عاملی بزرگی روی مشخصه ی پنهان اول هستند، این مشخصه عامل برگ نام گذاری شد. نشانگرهای محیط تنه و قطر تنه دارای بارهای

عاملی بزرگی روی مشخصه‌ی پنهان دوم بودند. این مشخصه عامل ابعاد تنه نام‌گذاری شد. نشانگرهای طول خوشه میوه و عرض خوشه میوه دارای بارهای عاملی بزرگی روی مشخصه‌ی پنهان سوم بودند این مشخصه عامل حجم خوشه نامیده شد. به همین ترتیب چون نشانگرهای وزن خوشه میوه تر و تعداد پسته‌های خوشه دارای بارهای عاملی بزرگی روی عامل چهارم بودند این عامل وزن خوشه نامیده شد. برای بررسی امکان تعیین رابطه‌ی بین چهار مشخصه‌ی پنهان و ۱۱ نشانگر مربوط از نمودار مسیر شکل ۲ استفاده شد. خطی بودن روابط در این مدل ساختاری به عنوان یک تقریب از روابط حقیقی، یک مفهوم اصیل و متعارف است.

پیکان‌های یک طرفه روابط علی بین مشخصه‌های پنهان و نشانگرها را نشان می‌دهند. بین مشخصه‌های پنهان هیچ‌گونه رابطه‌ی علی وجود ندارد و فقط بنا به ماهیت مشخصه‌ها و روش تعیین آن‌ها (روش پرومکس) بین آن‌ها روابط همبستگی یا کوواریانس وجود دارد. پیکان‌های قوسی دو طرفه این روابط را بین مشخصه‌های پنهان نشان می‌دهند. در این مدل‌ها مشخصه‌های پنهان را با دایره یا بیضی و مشخصه‌های آشکار یا نشانگرها را با مستطیل یا مربع نمایش می‌دهند. نشانگرها یا همان مشخصه‌های کمی مورد بررسی در حقیقت آثار یا معرف‌های مشخصه‌های پنهان هستند که قابل مشاهده و اندازه‌گیری می‌باشند. چون در تحلیل عاملی اکتشافی، تعیین عامل‌ها یا مشخصه‌های پنهان هدف است، بنابراین بایستی مشخصه‌های آشکار که قابل اندازه‌گیری هستند به صورت کامل و با دقت اندازه‌گیری شوند.

نظر به ضرورت تعیین مقیاس اندازه‌گیری مشخصه‌های پنهان، ضرایب مسیر مربوط به این مشخصه‌ها با برخی از نشانگرها، یک در نظر گرفته شد. همان‌طور که روی نمودار مسیر (شکل ۲) و برنامه‌ی تدوین شده ملاحظه می‌شود ضرایب مسیرهای مربوط به:

- ۱- عامل برگ با نشانگرهای طول برگ، عرض برگ و مساحت برگ.
- ۲- عامل ابعاد تنه با نشانگر محیط تنه.
- ۳- عامل حجم خوشه با نشانگرهای طول و عرض خوشه میوه.
- ۴- عامل وزن خوشه با نشانگر وزن خوشه میوه تر.

یک تعیین شد. e_1, e_2, \dots, e_{11} خطای اندازه‌گیری مربوط به هر نشانگر و $e_{12}, e_{13}, e_{14}, e_{15}$ مربوط به مشخصه‌های پنهان هستند. برنامه شماره‌ی ۱ برای تحلیل مدل شکل ۲ با استفاده از زبان سیمپلیس نرم‌افزار لیزرل تدوین شد. شکل ۳ خروجی لیزرل را مربوط به اجرای این برنامه نشان می‌دهد.

در این مدل تعداد ۲۵ پارامتر آزاد موجود است که توسط لیزرل برآورد شده‌اند و مقادیر آن‌ها روی مسیرهای مدل درج شده است. تعداد درآیه‌های معلوم و غیر تکراری ماتریس واریانس-کوواریانس ۱۱ مشخصه‌ی مدل هم

است. بنابراین درجه آزادی مدل $66 - 25 = 41$ است. لیزرل امیدریاضی متغیر تصادفی مربع کای $\frac{11(11+1)}{2} = 66$

برای درجه آزادی ۴۱ را $107/42$ و احتمال نظیر را $0/0000$ معین کرده است. مقدار احتمال تعیین شده نشان دهنده‌ی عدم شایستگی برازش مدل است. به این مفهوم که امکان بیان روابط بین مشخصه‌های پنهان تعیین شده و نشانگرهای مربوط در قالب مدل خطی تعیین شده وجود ندارد. یکی از دلایل رد شایستگی برازش مدل می‌تواند ناشی از کم بودن تعداد نشانگرهای سه مشخصه‌ی پنهان عامل ابعاد تنه، عامل حجم خوشه و عامل وزن خوشه باشد، چنان‌چه ملاحظه شد برای هر یک از این مشخصه‌ها کمتر از سه نشانگر تعریف شد. در اینجا بنا به عدم ضرورت از بیان کمیت‌های

مربوط به آزمون ضرایب مسیر و سایر شاخص‌های شایستگی برازش، موجود در خروجی لیزرل، خودداری شد. قابل ذکر است که اجرای پیشنهادهای خروجی برنامه نیز برای کاهش مقدار متغیر مربع کای و افزایش احتمال نظیر، بنا به ماهیت مشخصه‌های پنهان و نشانگرهای آن‌ها امکان‌پذیر نبود.

به این ترتیب در این بررسی با استفاده از روش تحلیل عاملی اکتشافی تعداد ۱۱ مشخصه‌ی آشکار در چهار مشخصه‌ی پنهان شامل: عامل برگ، عامل ابعاد تنه، عامل حجم خوشه و عامل وزن خوشه خلاصه شدند. این چهار مشخصه پنهان روی هم ۷۴ درصد واریانس مشخصه‌های آشکار را توضیح می‌دهند. بنابراین توجه به این مشخصه‌ها و نشانگرهای مربوط در برنامه‌های به نژادی اهمیت ویژه‌ای دارد. هر چند که به علت عدم شایستگی برازش مدل امکان بیان روابط بین این چهار مشخصه با ۱۱ نشانگر فراهم نشد. در انتها پیشنهاد می‌شود مشابه این مطالعه درباره‌ی مشخصه‌های ظاهری ارقام پسته سایر استان‌ها انجام شود. بدیهی است مقایسه نتایج می‌تواند مفید باشد.

منابع

اسماعیل‌پور، علی. ۱۳۷۷. بررسی، شناسایی و جمع‌آوری ارقام پسته. گزارش نهایی مؤسسه تحقیقات پسته کشور. ۲۴. اعلمی، علی. ۱۳۷۵. کاربرد ایزوآنزیم‌ها به منظور بررسی تنوع ژنتیکی ارقام پسته ایرانی. پایان‌نامه کارشناسی ارشد، دانشکده کشاورزی دانشگاه تربیت مدرس. ۹۱.

تاج‌آبادی‌پور، علی و محمد صانعی شریعت‌پناهی. ۱۳۷۶. شناسایی ارقام پسته ایران. پایان‌نامه کارشناسی ارشد. دانشکده کشاورزی دانشگاه تهران.

مرتضوی، سید عبدالمجید. ۱۳۷۶. گزارش پژوهشی مؤسسه تحقیقات پسته کشور. ایستگاه تحقیقات پسته دامغان. ۱۰.

- Alexander, D. M. 1984. Pistachio hybridization. *AUS. Hort.* 82(8): 54-58.
- Gerbing, D. W. and J.G. Hamilton (1996). Validity of Exploratory Factor Analysis. *Structural Equation Modeling*, 3. 62-72.
- IPGRI. 1996. Descriptors for Pistachio (*Pistacia Vera L.*)
- Maggs, D. H. 1981. Pistachio variety SIRORA. *Journal of the Australian Institute of Agricultural Science*.
- Mueller, R. O. (1996). *Basic Principles of Structural Equation Modeling: An Introduction to LISREL and EQS*. Secaucus, NJ: Springer.
- Parfit, D. E. 1999. Pistachio Cultivars improvement. 1999. California Pistachio Industry Annual Report. 1999-2000. P. 110-116.
- Stevens. J. (1996). *Applied Multivariate Statistics for the Ecological Science* (2nd ed.).
- Thurstone, L. L. (1935). *The Vectors of Mind: Multiple Factor Analysis for the Isolation of Primary Traits*. Chicago. Illinois: University of Chicago Press.
- Vargas, F. J., M. R. Romero, F. Monastera, A. Mendes Gaspar and D. Rouskas. 1998. Selection of Pistachio varieties adapted to the northern mediteranian region.

A Survey of Correlations between Morphological Traits in Pistachio by Exploratory Factor Analysis

Abstract

Despite high diversity Pistachio cultivars in Iran particularly in Kerman Province, no survey has been done about the nature of correlations between traits. In design and executing classic plant breeding programs based on selecting superior cultivars, depend on simple correlations between traits has no scientific explanation, because existing simple correlations between traits have no clear interpretation. Though, existence of these correlations can explain the existence of special relationships and common factors between traits. In current study, exploratory factor analysis was used for exploration and determination common factors or latent traits which cause correlation between observed traits. Therefore, by use of principle components analysis and promax rotation, four latent traits which state the most variabilities of studied traits were extracted and named. These four traits altogether explained 73/543 percent of variabilities of studied traits. Because goodness of fit of factor analysis model with $df = 41$ wasn't confirmed ($\chi^2 = 107/42$, $P\text{-value} = 0/0000$), the possibility of expression the relationships of four latent traits with observed traits in the form of a linear model as an approximation of actual relationships wasn't achieved.