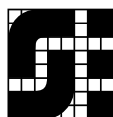


استفاده از روش‌های استنباط و پیش‌گویی بیزی برای پیش‌گویی نرخ بیکاری ایران (سال ۱۳۹۵)

اشکان شباک (مجری)

تابان باغ‌فلکی



گروه پژوهشی آمار

گروه پژوهشی پردازش داده‌ها و اطلاع‌رسانی

زمستان ۱۳۹۶

به نام خداوند جان و خرد

پیش‌گفتار

دشواری روزافزون گردآوری اطلاعات به روش‌های سنتی به دلیل پیچیدگی‌های امروزی جوامع آماری، نیاز به مطالعه برای دگرگونی و یا بهنگام سازی روش‌های آماری را ضروری کرده است. استفاده از دیگر منابع داده‌ها و مدل‌سازی، از روش‌هایی هستند که می‌توانند به عنوان جایگزین روش‌های آماری بکار رفته و یا به افزایش دقت برآوردها و استنباط‌های ناشی از آماری‌های سنتی یاری رسانند. منابع اطلاعاتی ناشی از داده‌های پیشین (مانند آماری‌های گذشته و یا داده‌های ثبت شده) همواره یکی از مهمترین منابع اطلاعاتی برای این منظور می‌باشند. بنابراین به نظر می‌رسد اگر بتوان با بکارگیری روش‌های استنباط بیزی و مدل سازی، از اطلاعات این منابع داده‌ها در پیش‌گویی یا بهنگام‌سازی برآوردهای موردنظر بهره جست، گامی بزرگ در جهت کاهش هزینه و خطای آماری‌ها به ویژه در آماری‌های رسمی، به شمار خواهد رفت.

در این طرح پژوهشی کوشیده می‌شود تا ضمن مطالعه‌ی مفاهیم و اصول نظریه‌ی بیزی، با مرور منابع موجود و بررسی آخرین پژوهش‌ها و اقدامات انجام شده درباره‌ی بکارگیری روش‌های استنباط بیزی، به امکان‌سنجی استفاده از این روش‌ها در آماری‌های مراکز آمار رسمی و به ویژه انجام پیش‌گویی‌های مورد نظر و هم‌پیوند با آماری‌های اشاره شده بپردازد. آنچه در این پژوهش دنبال می‌شود ضمن بررسی و شناساندن مفهوم و روش‌های موجود در آمار بیزی، پاسخ به این پرسش اصلی است که به‌طور کلی آیا می‌توان از توانایی‌های نهفته در روش‌های بیزی در پردازش و تحلیل اطلاعات گذشته، برای انجام پیش‌گویی آماری و افزایش دقت آماری‌ها در آمار رسمی بهره جست.

با توجه به مهم بودن موضوع اشتغال و بیکاری در عرصه اجتماعی و اقتصادی کشور و تاثیر پذیرفتن بسیاری از تصمیم‌ها و سیاست‌های کشور از این دو مقوله در این پژوهش، پیش‌گویی نرخ بیکاری بر اساس داده‌های رسمی به دست آمده از آماری نیروی کار مرکز آمار ایران در دستور کار قرار گرفته است. بنابراین با توجه به این‌که یکی از کاربردهای مهم روش‌های بیزی انجام پیش‌گویی است و از آن‌جایی که بر اساس مطالعات انجام شده تاکنون در کشور، پژوهشی در زمینه برآورد یا پیش‌گویی نرخ بیکاری به روش بیزی انجام نشده است، این تحقیق بنا دارد تا این کار را انجام داده و به‌عنوان اولین مطالعه عملی در این زمینه گشایش‌گر باشد. ضمن این‌که از نتایج این‌گونه پیش‌گویی‌ها می‌توان در راستی آزمایی نتایج آماری‌ها نیز بهره برد.

در پایان لازم است علاوه بر مجریان این طرح از جناب آقای نریمان یوسفی و سرکار خانم دکتر لیدا کلهری که با داوری و ارایه نظرهای ارزشمند در اصلاح و بهتر کردن هرچه بیشتر این طرح موثر بودند سپاسگزاری شود. همچنین از جناب آقای حامد لروند و سرکار خانم نجمه ناظریان که در تدوین و ویراستاری این طرح همکاری کردند نیز قدردانی می‌شود.

گروه پژوهشی پردازش داده‌ها و اطلاع‌رسانی

ژوئشکده‌ی آمار

فهرست مطالب

عنوان	شماره صفحه
۱. کلیات تحقیق	۱
۱-۱- مقدمه	۱
۲-۱- طرح مسأله	۳
۳-۱- اهمیت و ضرورت پژوهش	۵
۴-۱- اهداف تفصیلی پژوهش	۶
۵-۱- تعاریف و مفاهیم کلی مورد نیاز	۷
۲. مروری بر ادبیات پژوهش	۱۳
۱-۲- مقدمه	۱۳
۲-۲- آمارگیری از نیروی کار در ایران	۱۳
۳-۲- برآورد نرخ بیکاری	۱۷
۴-۲- پیش‌گویی نرخ بیکاری به روش بیزی	۲۱
۵-۲- سایر کاربردهای روش‌های بیزی در پیش‌گویی	۲۵
۳. استنباط بیزی	۲۹
۱-۳- مقدمه	۲۹
۲-۳- مدل‌بندی آماری	۲۹
۳-۳- آمار بیزی و آمار کلاسیک (فراوانی‌گرا)	۳۰
۴-۳- استنباط آماری بیزی	۳۲
۵-۳- گزینش توزیع‌های پیشینی در استنباط بیزی	۳۴
۶-۳- انتگرال‌گیری مونت‌کارلویی	۳۷
۷-۳- روش نمونه‌گیری رد-پذیرش	۳۷
۸-۳- الگوریتم متروپولیس-هستینگس	۳۸
۹-۳- نمونه‌گیری گیبز	۳۹
۱۰-۳- مزایا و معایب تحلیل بیزی با استفاده از نمونه‌گیری از توزیع پسینی	۳۹
۱۱-۳- عامل بیزی	۴۰
۴. تحلیل مدل‌های رگرسیونی طولی با استفاده از استنباط بیزی	۴۳

۴۳	۱-۴ - مقدمه
۴۳	۲-۴ - رگرسیون خطی ساده‌ی بیزی
۴۶	۳-۴ - مدل‌های بیزی برای تحلیل داده‌های طولی
۴۷	۱-۳-۴ - مدل اثرهای تصادفی برای تحلیل داده‌های طولی از دیدگاه بیزی
۴۹	۲-۳-۴ - مدل‌های حاشیه‌ای برای تحلیل داده‌های طولی از دیدگاه بیزی
۵۳	۳-۳-۴ - مدل‌های انتقالی برای تحلیل داده‌های طولی از دیدگاه بیزی
۵۵	۵. روش‌شناسی تحلیل بیزی نرخ بیکاری ایران (سال‌های ۱۳۸۴-۹۴)
۵۵	۱-۵ - مقدمه
۵۶	۲-۵ - تحلیل ابتدایی
۵۹	۳-۵ - تحلیل داده‌ها با استفاده از مدل حاشیه‌ای با ساختار کوواریانس $AR(1)$
۶۲	۴-۵ - تحلیل داده‌ها با استفاده از مدل حاشیه‌ای با ساختار کوواریانس $ARMA(1,1)$
۶۳	۵-۵ - تحلیل داده‌ها با استفاده از مدل حاشیه‌ای با ساختار کوواریانس $MA(1)$
۶۵	۶-۵ - تحلیل داده‌ها با استفاده از مدل اثرهای تصادفی
۶۶	۷-۵ - تحلیل داده‌های نرخ بیکاری با استفاده از مدل انتقالی مرتبه اول
۶۸	۸-۵ - تحلیل داده‌های نرخ بیکاری با استفاده از مدل انتقالی مرتبه دوم
۷۵	۶. نتیجه‌گیری و پیشنهادها
۷۵	۱-۶ - مقدمه
۷۵	۲-۶ - پیش‌گویی نرخ بیکاری برای سال ۱۳۹۵
۸۵	مرجع‌ها

کلیات تحقیق

۱-۱- مقدمه

با توجه به پیچیده‌تر شدن جوامع بشری، به دست آوردن آمار و اطلاعات از این جوامع نیز سخت‌تر و دشوارتر شده است. از این روی روش‌های گردآوری اطلاعات هرروز نیازمند دگرگونی و بهنگام سازی هستند. خوشبختانه پیشرفت فن‌آوری و دانش آمار ابزار مناسبی را در اختیار کارشناسان و کنش‌گران این حوزه گذاشته است تا با بهره‌گیری از آن‌ها بتوانند شیوه‌های نوینی را برای بهتر کردن روش انجام کار و بالا بردن دقت و کیفیت آمار گردآوری شده بنا نهند و یا با بهنگام سازی شیوه‌های سنتی کارایی آن‌ها را افزایش دهند. یکی از مهمترین مسایلی که همواره در روش‌های گردآوری آمار و اطلاعات وجود داشته است، بهینه‌گی هزینه‌ی انجام آمارگیری‌ها در برابر دقت برآوردها بوده است. اهمیت دقت و کیفیت برآوردهای ناشی از آمارگیری‌ها و نیز هزینه‌ی انجام آمارگیری‌ها باعث شده است که در طول زمان و به ویژه در دهه‌های اخیر دگرگونی‌های فراوانی در شیوه‌ی انجام آمارگیری‌ها و فنون به‌کاررفته در آن‌ها صورت پذیرد. یکی از برجسته‌ترین نکته‌هایی که در این زمینه به ذهن متبادر می‌شود بهره‌گیری از اطلاعات پیشین در برآوردها و استنباط‌هایی است که در آمارگیری‌ها انجام می‌شوند. در واقع استفاده از اطلاعات اضافی (کمکی) در علم آمار، پیشینه داشته و فراوان به‌کار رفته است. درعین حال، اگر بتوان از داده‌ها و اطلاعاتی که از گذشته در خصوص جامعه آمارگیری و واحدهای آماری آن موجود است، به طور نظام‌مند در پیش بینی مشاهدات جامعه آمارگیری بهره جست و یا آن‌ها را در بهنگام‌سازی و افزایش کیفیت برآوردهای انجام شده به‌کار گرفت گامی بزرگ در جهت بهینه‌سازی آمارگیری‌ها، انجام شده است. هنگامی که استفاده نظام‌مند و هدفمند از داده‌ها و اطلاعات گذشته مطرح می‌شود روش‌های بیزی به عنوان رویکردی مشخص و معین، توسط آمارشناسان به‌کارگرفته می‌شوند. در این فصل ضمن بیان مساله و ضرورت اجرای طرح به معرفی کلیات دیگر تحقیق از قبیل تعاریف و مفاهیم کلی مورد نیاز می‌پردازیم.

۱-۲- طرح مساله

همان‌گونه که در بالا گفته شد، کیفیت داده‌ها در آمارگیری‌ها، اندازه‌ی دقت برآوردها و نیز اعتبار استنباط‌های پس از آمارگیری از مهمترین و اصلی‌ترین موضوعاتی هستند که دغدغه‌ی دست اندرکاران آمارگیری و دیگر کارشناسان و آمارشناسان را تشکیل می‌دهند. در این میان هزینه یک آمارگیری همواره متغیری در برابر دقت و کیفیت آن بوده است. از سوی دیگر مشکل بی‌پاسخی، به طور روزافزونی در آمارگیری‌ها افزایش می‌یابد که هم منجر به افزایش هزینه و نیز

خطای آمارگیری می‌شود. خطای پوشش نیز از دیگر عواملی است که به دلایل گوناگون و گاه ناگزیر (مانند شرایط جغرافیایی) رخ می‌دهد که همراه با افزایش هزینه، دقت آمارگیری را به شدت کاهش می‌دهد. محدود بودن پاسخگویان و یا جامعه‌ی آمارگیری نیز مشکل دیگری است که (به ویژه در آمارگیری‌های جمعیتی) آمارشناسان و آمارگیران اغلب با آن روبرو هستند. از همین رو است که به ویژه در سال‌های اخیر لزوم تغییر در روش‌های آمارگیری و یا جایگزینی آن‌ها مورد توجه آمارشناسان بوده است. به‌عنوان مثال توجه به آمارهای ثبتي و به‌کارگیری داده‌های ثبتي در آمارگیری‌ها (آمارگیری‌های ثبتي مبنا) از نخستین و البته مهمترین روش‌هایی است که مورد توجه آمارشناسان بوده و به ویژه در سال‌های اخیر (همزمان با پیشرفت فناوری‌های رایان‌های و ارتباطی) کاربرد و اهمیت بیشتری پیدا کرده است.

از دیگر روش‌هایی که برای کاهش خطای آمارگیری به ویژه در آمارگیری‌هایی نمونه‌ای که در ناحیه‌های کوچک انجام می‌شوند مورد نظر آمارشناسان قرار گرفته است آمارگیری‌های مدل-مبنا^۱ می‌باشد. البته این روش به جز کاربرد در موارد معدود مانند آمارگیری‌های ناحیه‌های کوچک، هنوز در دیگر موارد گسترش نیافته و در حال نظریه‌پردازی است. (در حال حاضر بیشتر روش‌های آمارگیری نمونه‌ای طرح-مبنا^۲، مدل یار-مبنا^۳ هستند.) برای توضیح بیشتر باید گفت که به طور کلی آمارگیری نمونه‌ای برای بدست آوردن اطلاعات آماری از پارامترهای یک جامعه با مطالعه‌ی موردی نمونه‌ای کوچکتر از آن جامعه طراحی و انجام می‌شود. استنباط آماری برای برآورد کردن و یا پیش‌بینی پارامترهای مورد نظر در جامعه نیز معمولاً به یکی از رویکردهای «طرح-مبنا، مدل یار-مبنا و یا مدل-مبنا» صورت می‌گیرد. روش طرح-مبنا، همان روش سنتی و معمول است که بر اساس نظریه نمونه‌گیری و بر پایه‌ی توزیع نمونه‌ای برآوردگرها، ناشی از طرح نمونه‌گیری، پایه‌ریزی شده است. روش مدل-مبنا با تأکید کمتر بر توزیع نمونه‌ای برآوردگرها، تلاش دارد با استفاده از گمانه‌زنی یک مدل آماری به برآورد پارامترهای مورد نظر و یا انجام پیش‌بینی‌های لازم بپردازد. در روش مدل یار-مبنا نیز که بیشتر به روش طرح-مبنا نزدیک است، استنباط‌های آماری با استفاده از اطلاعات کمکی (صفات یا متغیرهای کمکی) دقیق‌تر می‌شوند. معمولاً در این روش برآوردها و پیش‌گویی‌ها با به‌کارگیری یک مدل رگرسیونی و یا سری زمانی که بر پایه ارتباط میان پارامترهای مورد نظر و متغیرهای کمکی شکل گرفته است به دست می‌آیند. این برآوردها سپس با توجه به طرح نمونه‌گیری و ویژگی‌های توزیع نمونه‌ای تعدیل شده و برای استنباط نهایی درباره پارامترها به‌کار می‌روند. به دلایل ویژگی‌های آماری خوبی که این‌گونه استنباط‌ها (مدل مبنا) دارند و افزایش دقت برآوردها با بهره‌گیری از صفات کمکی، کاربرد گسترده‌ای در انجام آمارگیری‌های نمونه‌ای دارند. به‌ویژه در مواردی که به دلایل هزینه و یا دسترسی به نمونه‌ها امکان انجام نمونه‌گیری به آسانی مهیا نباشد.

یکی دیگر از شیوه‌هایی که می‌تواند برای غلبه بر مشکلات معمول در آمارگیری‌های سنتی و یا کاهش اثرهای آن‌ها، مطرح باشد انجام پیش‌گویی^۴ و یا پیش‌بینی^۵ آماری در آمارگیری‌ها است. پیش‌گویی/پیش‌بینی یکی از بخش‌های مورد توجه در استنباط آماری است که در راستای روش‌های مدل مبنا انجام شده و همواره مورد توجه آمارشناسان بوده است. به طور کلی در استنباط آماری پیش‌گویی، رویکردی تعریف می‌شود که در آن مشاهده‌های آینده بر پایه‌ی مشاهده‌های اندازه‌گیری شده در گذشته، محاسبه (پیش‌گویی) می‌شوند. در آمار کلاسیک همیشه برای انجام پیش‌گویی

¹ Model-based

² Designed-based

³ Model- assisted based

⁴ Prediction

⁵ Forecast

مدل‌سازی شده است. آنالیز مدل‌های رگرسیونی و مدل‌های سری زمانی برای انجام پیش‌گویی به‌کار می‌روند. زمانی که پیش‌گویی مشاهده‌ها بر پایه‌ی اطلاعات دقیق از توزیع پیشین پارامترهای مورد نظر جامعه آماری انجام شود در واقع پیش‌گویی بیزی انجام شده است. در روش‌های استنباط بیزی^۶، پیش‌گویی چه در قالب به دست آوردن توزیع پسین پیش‌گو^۷ برای مشاهدات آینده و چه در قالب مدل‌سازی بیزی (به دست آوردن برآوردگرهای بیزی و گنجاندن آن‌ها در مدل‌های مفروض) می‌تواند انجام شود.

۱-۳- اهمیت و ضرورت پژوهش

همان‌گونه که اشاره شد دشواری روزافزون گردآوری اطلاعات به روش‌های سنتی به دلیل پیچیدگی‌های امروزی جوامع آماری، نیاز به مطالعه برای دگرگونی و یا بهنگام‌سازی روش‌های آماری را ضروری کرده است. استفاده از دیگر منابع داده‌ها و مدل‌سازی، از روش‌هایی هستند که می‌توانند به عنوان جایگزین روش‌های آماری به‌کار رفته و یا به افزایش دقت برآوردها و استنباط‌های ناشی از آماری‌های سنتی یاری رسانند. منابع اطلاعاتی ناشی از داده‌های پیشین (مانند آماری‌های گذشته و یا داده‌های ثبت شده) همواره یکی از مهمترین منابع اطلاعاتی برای این منظور می‌باشند. بنابراین به نظر می‌رسد اگر بتوان با به‌کارگیری روش‌های استنباط بیزی و مدل‌سازی، از اطلاعات این منابع داده‌ها در پیش‌گویی یا بهنگام‌سازی برآوردهای موردنظر بهره جست، گامی بزرگ در جهت کاهش هزینه و خطای آماری‌ها به ویژه در آماری‌های نمونه‌ای، به شمار خواهد رفت.

بنا بر آن‌چه گفته شد به‌طور خلاصه می‌توان جمع‌بندی کرد که در این طرح قصد داریم تا ضمن مطالعه‌ی مفاهیم و اصول نظریه‌ی بیز، با مرور منابع موجود و بررسی آخرین پژوهش‌ها و اقدامات انجام شده درباره‌ی به‌کارگیری روش‌های استنباط بیزی، به امکان‌سنجی استفاده از این روش‌ها در آماری‌های مراکز آمار رسمی و به ویژه انجام پیش‌گویی‌های مورد نظر و هم‌پیوند با آماری‌های اشاره شده بپردازد. آنچه در این پژوهش دنبال می‌شود ضمن بررسی و شناساندن مفهوم و روش‌های موجود در آمار بیز، پاسخ به این پرسش اصلی است که به‌طور کلی آیا می‌توان از توانایی‌های نهفته در روش‌های بیزی در پردازش و تحلیل اطلاعات گذشته، برای انجام پیش‌گویی آماری و افزایش دقت آماری‌ها در آمار رسمی بهره جست. در این تحقیق برای پاسخ به این پرسش، پیش‌گویی نرخ بیکاری را بر اساس داده‌های رسمی به دست آمده از آماری‌های نیروی کار مرکز آمار در دستور کار قرار داده‌ایم.

استفاده از اطلاعات کمکی در آماری‌ها پیشینه‌ی طولانی دارد و امروز نیز فراوان در آماری‌های نمونه‌ای به ویژه در نمونه‌گیری‌های مدل-یاری‌منا به‌کار گرفته می‌شود. اما بهره‌گیری از داده‌ها و اطلاعات پیشین به صورت نظام‌مند که برای تعیین توزیع داده‌های نمونه‌گیری شده به‌کار می‌روند گرچه بسیار جالب به نظر می‌رسد پیشینه کمتری دارد. بدون شک در صورتی که بتوان به کمک داده‌های پیشین تقریب مناسبی از توزیع داده‌ها و یا مدل‌های مفروض برآوردهای آن‌ها، ارائه داد می‌توان با انجام پیش‌گویی در مورد مشاهدات آینده، گامی بلند در جهت کاهش هزینه‌های آماری، افزایش دقت برآوردها (با روش بهنگام‌سازی اطلاعات) و حتی راست‌آزمایی برآوردهای انجام شده در آماری برداشت. به‌طور کلی انجام پیش‌بینی و یا پیش‌گویی به روش‌های بیزی پیشینه و تجربه عملی چندانی در دنیا ندارد و در

⁶ Bayesian inferences

⁷ Posterior predictive distribution

کشور ما نیز به تبع آن تاکنون مطالعه‌ای در خصوص به‌کارگیری این روش برای استنباط در آمار رسمی یافت نشده است. بنابراین این پژوهش می‌تواند برای ادامه مطالعات و پژوهش‌ها برای بهره‌گیری از روش‌های بیزی در تحلیل‌های آماری مدل مبنا و کاربرد عملی آن در آمار رسمی راه‌گشا باشد.

یک مسأله متداول در همه جوامع بشری، پایش و کنترل شاخص‌های مهم اقتصادی - اجتماعی است. اشتغال و بیکاری، از جمله موضوعات اساسی اقتصاد هر کشور است، به گونه‌ای که افزایش اشتغال و کاهش بیکاری، به عنوان یکی از شاخص‌های توسعه‌یافتگی جوامع تلقی می‌شود. نرخ بیکاری یکی از شاخص‌هایی است که برای ارزیابی شرایط اقتصادی کشور مورد استفاده قرار می‌گیرد. پایش و اعمال سیاست در خصوص کنترل یا ایجاد تغییر دلخواه در نرخ بیکاری از دغدغه‌های دولتمردان محسوب می‌شود. طرح آمارگیری نیروی کار با هدف برآورد شاخص‌های نیروی کار به صورت فصلی و سالانه و همچنین تغییرات آن در کل کشور، نقاط شهری، روستایی و تغییرات سالانه در استان‌ها با روش نمونه‌گیری چرخشی^۸ در فصول مختلف سال اجرا می‌شود.

بنابراین اگر بتوان الگویی برای پیش‌گویی نرخ بیکاری با استفاده از روش‌های بیزی ارائه داد، می‌توان حداقل در راست آزمایی داده‌های به دست آمده از آمارگیری نیروی کار، تصحیح و تعدیل داده‌های به دست آمده و یا حتی کاهش هزینه‌های اجرای آن با اجرا نکردن طرح در برخی سالها و یا فصول خاص دستاوردهای چشمگیری داشت. از آنجا که یکی از کاربردهای مهم روش‌های بیزی انجام پایش و برآورد است، اما بر اساس مطالعات انجام شده تاکنون در کشور، پژوهشی در زمینه برآورد و پیش‌بینی نرخ بیکاری به روش بیزی انجام نشده است که این پژوهش بنا دارد تا این کار را انجام داده و به‌عنوان اولین مطالعه عملی در این زمینه گشایش‌گر این راه باشد.

۱-۴- اهداف تفصیلی پژوهش

همان‌گونه که اشاره شد و از عنوان این طرح نیز برمی‌آید مهمترین هدف این پژوهش این است که به طور کلی به این پرسش پاسخ دهد که «آیا می‌توان با بهره‌گیری از داده‌ها و اطلاعاتی که از گذشته تاکنون در خصوص جامعه آمارگیری و واحدهای آماری آن موجود است، توسط روش‌های استنباطی بیزی به پیش‌گویی آماری در یک جامعه آمارگیری پرداخت؟» با توجه به این هدف کلی و با محوریت آن، انتظار می‌رود در پایان این پژوهش، بتوان به پرسش‌های تفصیلی‌تر زیر نیز پاسخ داد:

- مفاهیم، اصول کلی و روش‌های آمار استنباطی بیزی و پیش‌گویی بیزی چه هستند؟
- چه اقدامات و مطالعاتی توسط دیگر آمارشناسان و یا سازمان‌های آمار رسمی جهان در زمینه استفاده از روش‌های استنباطی بیزی در آمارگیری‌ها انجام شده است؟
- آیا به طور کلی پیش‌گویی‌های بیزی در آمارگیری‌های رسمی قابلیت به‌کارگیری دارند؟ و اگر پاسخ مثبت است چگونه می‌توان از این روش‌های استفاده کرد؟

⁸ Rotation Sampling

۱-۵- تعاریف و مفاهیم کلی مورد نیاز

در این بخش به شناساندن برخی مفاهیم کلی که در این طرح مطرح می‌شوند می‌پردازیم. شایان گفتن است که دیگر مفاهیم تخصصی و فنی مورد نیاز در بخش‌های در پیوند با همان مفاهیم تخصصی و فنی ارائه خواهند شد.

کار: آن دسته از فعالیت‌های اقتصادی (فکری یا بدنی) که به منظور کسب درآمد (نقد یا غیرنقدی) صورت پذیرد و هدف آن تولید کالا یا ارائه خدمت باشد، کار محسوب می‌شود.

آمارگیری مکرر^۹: بسیاری از آمارگیری‌های مهم در فواصل زمانی معین و بطور منظم اجرا می‌شوند. می‌توان از داده‌هایی که طی نمونه‌گیری مکرر از یک جامعه در دوره‌ای از زمان بدست می‌آید، به منظور برآورد پارامترهای مربوط به ویژگی‌های مورد مطالعه برای دوره زمانی جاری، برآورد تغییرات پارامترها از دوره قبل تا دوره زمانی جاری، و برآورد پارامترهایی مانند میانگین یا مجموع کل برای چند دوره زمانی استفاده کرد. همچنین می‌توان با تلفیق داده‌های دوره زمانی جاری و داده‌های قبلی، برآوردهای دوره‌های زمانی را تعدیل و اصلاح نمود. آمارگیری‌های مکرر علاوه بر اینکه از وضعیت جامعه مورد مطالعه در زمان جاری اطلاعاتی بدست می‌دهند، امکان شناخت عوامل موثر بر جامعه مورد مطالعه در طی زمان را نیز فراهم می‌کند.

در طراحی یک آمارگیری مکرر دو مسأله مطرح می‌شود:

(الف) تعداد دفعات نمونه‌گیری

(ب) الگوی ترکیب نمونه مربوط به هر زمان با نمونه دوره‌های زمانی قبلی

تعداد دفعات آمارگیری باید با در نظر گرفتن ضرورت دستیابی به برآوردهای دوره‌ای، میزان تغییرات پارامترهای مورد نظر در طی زمان، هزینه و مدت زمان لازم برای حصول برآوردهایی با دقت مطلوب انتخاب شود. الگوی ترکیب در نمونه‌گیری مکرر ممکن است به گونه‌ای باشد که در هر زمان:

(الف) نمونه شامل همان مجموعه واحدهایی شود که در دوره قبل نمونه‌گیری شده‌اند.

(ب) نمونه شامل مجموعه واحدهایی کاملاً متفاوت با واحدهای نمونه‌گیری شده در دوره قبل شود.

(پ) نمونه، مجموعه‌ای از واحدهای نمونه‌گیری شده در دوره قبلی و مجموعه‌ای از واحدهایی که در دوره قبل نمونه‌گیری نشده‌اند را شامل شود.

نمونه‌گیری مکرر با استفاده از الگوی سوم را نمونه‌گیری چرخشی گویند.

نمونه‌گیری چرخشی: در نمونه‌گیری چرخشی ممکن است برخی واحدها برای چند دوره زمانی متوالی در نمونه باقی بمانند و یا برخی واحدها پس از خروج از نمونه مجدداً در دوره‌های بعدی وارد نمونه شوند. منظور از چرخش نمونه جایگزین شدن همه یا تعدادی از واحدهای نمونه با واحدهای جدید از یک دوره به دوره دیگر است. در بسیاری از آمارگیری‌ها الگوی چرخش متقارن است به این معنی که مجموعه واحدهای جدید طی دوره‌های زمانی منظم به نمونه وارد می‌شوند و پس از ورود به نمونه هر مجموعه با تبعیت از

⁹ Repeated Survey

الگوی یکسان در نمونه باقی مانده و یا حذف می‌شود. با برقراری این سیستم میزان تداخل نمونه در دوره‌های متوالی و بین دو دوره با فاصله زمانی یکسان ثابت خواهد بود.

در بسیاری از آمارگیری‌ها از یک الگوی چرخش ساده استفاده می‌شود. نمونه شامل n زیر نمونه است در ابتدای هر دوره آمارگیری یک زیر نمونه جدید وارد نمونه می‌شود و هر زیر نمونه برای n بار (دوره) متوالی در نمونه باقی می‌ماند. میزان تداخل بین دوره‌ها با افزایش فاصله زمانی بین آن‌ها به‌طور خطی کاهش می‌یابد میزان تداخل برای دو نمونه با فاصله زمانی i ، تا زمانی که $i=n-1$ باشد برابر $\frac{n-1}{n}$ و برای $i>n$ برابر صفر است. نکته دیگری که باید به آن توجه داشت این است که در بسیاری موارد در یک نمونه‌گیری چند مرحله‌ای نمونه در سطح واحدهای مراحل بالاتر به‌کندی و در سطح واحدهای مراحل پایین‌تر با سرعت بیشتری چرخش می‌کند یعنی مثلاً در یک نمونه‌گیری دو مرحله‌ای نمونه در سطح PSU¹⁰ ها با سرعت کمتری نسبت به USU¹¹ ها تغییر می‌کند حتی در مواردی نمونه در سطح واحدهای مراحل بالاتر ثابت مانده و چرخش نمی‌کند (یکی از انواع روش‌های نمونه‌گیری، نمونه‌گیری خوشه‌ای است. در نمونه‌گیری خوشه‌ای، گروه‌ها یا خوشه‌هایی از اعضای از گروه‌هایی که بین اعضای آن‌ها نوعی عدم تجانس وجود دارد، برای مطالعه انتخاب می‌شوند. این نمونه‌گیری به دو شیوه تک مرحله‌ای و چند مرحله‌ای انجام می‌شود. در نمونه‌گیری خوشه‌ای تک مرحله‌ای، جامعه به خوشه‌ها یا گروه‌های مناسب تقسیم می‌شود و آزمودنی‌های نمونه، به صورت تصادفی از تعدادی خوشه‌ها یا گروه‌های مورد نیاز، انتخاب می‌شوند و همه‌ی عناصر از خوشه‌ای که به صورت تصادفی انتخاب شده است مورد بررسی قرار می‌گیرند. در نمونه‌گیری چند مرحله‌ای ابتدا از خوشه‌ها یا گروه‌های جامعه، نمونه‌ای اولیه به شیوه تصادفی استخراج می‌کنیم، که به آن واحد نمونه ای اولیه PSU می‌گوییم. سپس از بین این خوشه‌ها بار دیگر نمونه‌ای ثانویه SSU¹² تصادفی انتخاب می‌کنیم و این کار را تا زمانی که به مرحله نهایی این سطح برسیم انجام می‌دهیم. آنگاه در آخر نمونه‌ای خواهیم داشت که از هر واحدی، عضوی در آن وجود دارد، که به آن واحد نمونه‌ای نهایی USU می‌گوییم.) این امر بدلیل کاهش هزینه و دشواری تغییر واحدهای نمونه‌گیری مرحله اول (PSUها) و سایر واحدهای مراحل بالاتر است در سطح نهایی غالباً الگوی چرخش به جای خانوارها و افراد برای آدرس‌ها و یا سایر واحدهای ساختاری اعمال می‌شود، زیرا احتمال تغییر مکان خانوارها و افراد زیاد است و این امر موجب کاهش تداخل واقعی بین دوره‌ها می‌شود.

انتخاب یک الگوی خوب چرخشی به همبستگی بین جواب‌ها از یک واحد نمونه‌گیری شده در دو دوره زمانی متوالی بستگی دارد البته عوامل دیگری نیز هستند که با همین اهمیت هستند ولی به راحتی نمی‌توان آن‌ها را به صورت کمی بیان نمود. از میان دیگر عامل‌ها بار تحمیل شده به پاسخگوهایی که چند بار نمونه‌گیری می‌شوند می‌باشد.

¹⁰ Primary Sampling Unit

¹¹ Ultimate Sampling Unit

¹² Secondary Sampling Unit

از نقطه نظر اداری هزینه نمونه‌گیری از واحد، هزینه ای است که برای قرار گرفتن واحد در بار اول نمونه‌گیری پرداخته می‌شود. این هزینه‌ها شامل فهرست‌برداری، شناسایی و مکان‌یابی واحدها می‌باشد. مهم‌ترین شاخص‌هایی که در طرح نیروی کار محاسبه می‌شوند به شرح زیر می‌باشند:

در این طرح از بعضی اصطلاحات آماری مانند برآورد، پیش‌بینی، پیش‌گویی و ... نام برده‌ایم که در زیر آن‌ها را تعریف می‌کنیم و تفاوت آن‌ها را مشاهده خواهیم کرد.

برآورد^{۱۳}: برآورد یا تخمین به روشی گفته می‌شود که برای پیش‌بینی یا حدس تعداد اعضای یک گروه با استفاده از تقریب استفاده می‌شود. معمولاً زمانی به‌کار می‌رود که ورودی‌های بدست آمده کامل نباشند یا عدم قطعیت در تعداد آن‌ها باشد یا ثبات وضعیت وجود نداشته باشد. در علم آمار برآورد بر دو نوع است: برآورد کلاسیک و برآورد بیزی. در برآورد کلاسیک هدف این است که بر اساس مشاهدات یک متغیر تصادفی پارامتر مجهول را برآورد کنیم که به دو صورت انجام می‌گیرد: برآورد نقطه‌ای و برآورد فاصله‌ای. در برآورد نقطه‌ای بر اساس مشاهدات یک نقطه را برای پارامتر تخمین می‌کند اما در برآورد فاصله‌ای یک فاصله را برای پارامتر برآورد می‌کند. روش بیزی بر خلاف روش کلاسیک که پارامتر را به‌عنوان یک مقدار ثابت و مجهول در نظر می‌گیرند، پارامتر را به‌عنوان یک متغیر تصادفی در نظر می‌گیرد. در واقع در روش بیزی هدف این است که بر اساس مشاهدات متغیر تصادفی، متغیر تصادفی دیگری را تخمین کنیم.

پیش‌گویی^{۱۴}: معمولاً بر اساس تجربیات افراد استوار است و بسته به میزان تجربه و اطلاعات شخص از یک موضوع مشخص صحت و سقم پیش‌گویی متغیر می‌باشد هرچه میزان دانسته‌ها و تجربیات بیشتر باشد درستی برآورد روند یک موضوع در آینده بیشتر است بنابراین بهتر است از تجربیات افراد مختلف استفاده شود. معمولاً از پیش‌گویی برای برآوردهای کلی پارمترها در آینده استفاده می‌شود و یک موضوع مهم در استنباط آماری است و چه برای داده‌های مقطعی^{۱۵} و چه طولی^{۱۶} انجام می‌شود.

پیش‌بینی^{۱۷}: پیش‌بینی عبارتست از تجسم یک موقعیت در آینده براساس اطلاعات گذشته. در واقع در پیش‌بینی بر اساس معیارهای کمی از داده‌های به وقوع پیوسته در زمان‌های گذشته برای تخمین آینده استفاده می‌کنیم که به چنین داده‌هایی سری‌های زمانی می‌گویند به عبارت دیگر پیش‌بینی بر اساس معیارهای کمی در واقع قسمتی از مباحث سری‌های زمانی است. در این حالت پیش‌بینی برای یک زمان مشخص در آینده رخ می‌دهد.

با توجه به تعریف بالا گذاشتن تفاوت میان پیش‌گویی و پیش‌بینی کمی دشوار است. در روش‌های آماری تنها در زمانی که پیش‌گویی بر پایه روش‌های موسوم به سری زمانی انجام می‌شود به آن واژه پیش‌بینی اطلاق می‌شود. با این وجود می‌توان ادعا کرد که در هر حال پیش‌گویی مفهوم عام‌تری از پیش‌بینی است و پیش‌بینی زیرمجموعه آن قرار دارد.

¹³ Estimation

¹⁴ Prediction

¹⁵ cross-sectional data

¹⁶ longitudinal data

¹⁷ Forecasting

استنباط بی‌زی^{۱۸}: در روش‌های کلاسیک پارامتر θ را یک مقدار ثابت نامعلوم در نظر می‌گیرند و یک نمونه تصادفی X_1, \dots, X_n از جمعیت که دارای توزیع $f_\theta(x)$ بود جمع‌آوری کرده و براساس آن در مورد θ تصمیم‌گیری می‌شود. در روش بی‌زی θ را کمیتی در نظر می‌گیریم که خود یک متغیر تصادفی است و تغییرات آن توسط یک توزیع احتمال (توزیع پیشین^{۱۹}) بیان می‌گردد. سپس از جمعیت یک نمونه آماری جمع‌آوری می‌شود و بر اساس آن توزیع پیشین تصحیح می‌گردد. این توزیع تصحیح شده همان توزیع پسین^{۲۰} است که در بالا آن را تعریف کردیم. سپس بر اساس این توزیع تصمیم‌گیری انجام می‌شود. برای مثال فرض کنید قطعات تولیدی یک کارخانه دارای توزیع نمایی با میانگین θ ساعت می‌باشند و به علت فرسودگی دستگاه‌ها این میانگین θ نیز در سال تغییر می‌کند و تغییرات آن طبق یک توزیع احتمال $\pi(\theta)$ می‌باشد، $\pi(\theta)$ توزیع پیشین می‌باشد.

توزیع پیشین احتمال: در استنباط بی‌زی، توزیع پیشین احتمال، توزیع احتمالی است که قبل از مشاهده داده‌ها برای یک مقدار نامعلوم بیان می‌شود. این توزیع، براساس اعتقادات و تجربه‌های قبلی آزمایشگر و قبل از مشاهده داده‌ها تعیین می‌گردد. این مقدار نامعلوم، اغلب پارامتر^{۲۱} یا یک متغیری است که به‌طور مستقیم قابل مشاهده نیست. برای مثال، می‌توان قبل از پایان رأی‌گیری در انتخابات نسبت رأی‌دهندگان به یک نامزد خاص را با یک توزیع احتمال مدل‌بندی نمود.

توزیع پسین احتمال: توزیع پسین، توزیع احتمال یک مقدار نامعلوم به شرط مشاهده داده‌ها است. بر خلاف توزیع پیشین، در توزیع پسین بعد از مشاهده داده‌ها توزیع احتمال را برای پارامتر بیان می‌کنیم. عبارت «پسین»، به دلیل اینکه توزیع را بعد از در نظر گرفتن داده‌ها بیان می‌کنیم، به‌کار می‌رود و به صورت $\pi(\theta|X)$ نمایش می‌دهند، احتمال پارامتر θ به شرط مشاهده X است. روش به دست آوردن توزیع پسین به شرح زیر است:

فرض کنید θ دارای توزیع پیشین $\pi(\theta)$ باشد و $D = (D_1, \dots, D_n)$ یک نمونه تصادفی جمع‌آوری شده از جمعیتی باشد که توزیع آن به θ بستگی دارد. بر اساس این نمونه تصادفی در توزیع پیشین $\pi(\theta)$ بازنگری می‌کنیم و چگالی پسین θ را به صورت زیر بدست می‌آوریم:

$$\pi(\theta|D) = \frac{f(d, \theta)}{m(d)}$$

که در آن $f(d, \theta)$ چگالی توام D و θ و $m(d)$ چگالی کناری D است بنابراین

$$f(d, \theta) = \pi(\theta)f(d|\theta)$$

¹⁸ Bayesian Inference

¹⁹ Prior distribution

²⁰ Posterior distribution

²¹ Parameter

$$m(d) = \int f(d, \theta) d\theta$$

پس

$$\pi(\theta|D) = \frac{\pi(\theta)f(d|\theta)}{m(d)} = c(d)\pi(\theta)f(d|\theta) \propto \pi(\theta)f(d|\theta) = \pi(\theta)L(\theta) \quad (1-0)$$

یعنی برای محاسبه توزیع احتمال پسین θ کفایت حاصل ضرب $\pi(\theta)$ و $L(\theta)$ (تابع درست‌نمایی) را بدست آورده و آنرا به یک توزیع احتمال تبدیل کنیم.

شایان گفتن است که با توجه به توزیع پسین به دست آمده در رابطه (1-0)، می‌توان در مورد پارامتر تصمیم‌گیری کرد که به آن تصمیم بیزی^{۲۲} گفته می‌شود.

پارامتر: کمیت عددی و معمولاً نامعلوم می‌باشد که مشخصه‌ای معین در مورد کل جامعه را شرح می‌دهد مانند میانگین یا مُد جامعه. پارامترها چون مقدارشان معمولاً نامعلوم است اغلب برآورد می‌شوند؛ به خصوص زمانی که جامعه به اندازه کافی بزرگ باشد حصول اندازه‌های مربوط به کل جامعه، غیر ممکن یا غیر عملی است (به طور قراردادی پارامترها را حروف یونانی نمایش می‌دهند).

تابع درست‌نمایی^{۲۳}: در آمار، تابع درست‌نمایی (که اغلب برای سادگی درست‌نمایی خوانده می‌شود) تابعی از پارامترهای مدل آماری است که نقش کلیدی در آمار استنباطی ایفا می‌کند. تابع درست‌نمایی برابر است با احتمال آنکه به ازای مقادیری معین برای تعدادی پارامترهای آماری، نتایج مشاهده شده حاصل شود. به زبان ریاضی، اگر X مجموعه‌ای از داده‌های مشاهده شده و Θ مجموعه‌ای از پارامترها باشد، آنگاه احتمال شرطی $P(X|\Theta)$ را می‌توان مشاهده $L(X|\Theta)$ ، احتمال درست‌نمایی Θ براساس X دانست. حال اگر X ثابت نگه داشته شود و Θ تغییر داده شود، بوضوح می‌توان دید که $L(X|\Theta)$ تابعی از Θ است.

توزیع پسین پیشگو^{۲۴}: در آمار بیزی، توزیع پسین پیشگو، توزیع مقادیر غیر مشاهده به شرط مقادیر مشاهده شده است. فرض کنید به شرط مجموعه مشاهدات $X = \{X_1, \dots, X_n\}$ ، بخواهیم توزیع پسین مقدار جدید \tilde{X} را پیشگویی کنیم، برای این کار از رابطه زیر استفاده می‌کنیم:

$$p(\tilde{X}|X) = \int p(\tilde{X}|\theta, X)P(\theta|X, \alpha)d\theta$$

²² Bayesian decision

²³ Likelihood function

²⁴ Posterior predictive distribution

مروری بر ادبیات پژوهش

۲-۱- مقدمه

در این فصل تلاش خواهد شد تا با مرور ادبیات پژوهش مورد نظر به ارائه‌ی مبانی نظری و علمی هم‌پیوند با این تحقیق پرداخته شود. این مرور ادبیات در دو بخش انجام شده است. از آنجایی که در این تحقیق در عمل با بهره‌گیری از داده‌های آمارگیری نیروی کار، روش‌های بیزی برای پیش‌گویی نرخ بیکاری استفاده شده است، ابتدا و در بخش نخست به مرور تاریخچه و مبانی پایه‌ای این آمارگیری مهم در مرکز آمار ایران پرداخته خواهد شد. سپس و در ادامه بحث به بررسی مطالعات و پژوهش‌هایی که در خصوص برآورد و پیش‌بینی نیروی کار و نیز استفاده از روش‌های بیزی برای پیش‌بینی شده است، به اجمال اشاره خواهد شد.

۲-۲- آمارگیری از نیروی کار در ایران

آمار و اطلاعات مربوط به نیروی انسانی، برای نخستین بار در سال ۱۳۱۸ توسط اداره کل آمار و ثبت احوال، در برخی از شهرهای مهم کشور گردآوری شد. در سال ۱۳۳۵، ضمن اجرای اولین سرشماری عمومی نفوس که توسط اداره کل آمار عمومی اجرا شد، اطلاعاتی چند در این زمینه به دست آمد. سپس با اجرای چند طرح آماری در زمینه‌های کشاورزی، صنعت و جمعیت، در زمره آمار جمع‌آوری شده، اطلاعات مربوط به نیروی انسانی نیز متناسب با موضوع طرح، مورد توجه قرار گرفت. در سال‌های بعد، مرکز آمار ایران ضمن اجرای سرشماری‌های عمومی نفوس و مسکن سال‌های ۱۳۴۵، ۱۳۵۵، ۱۳۶۵، ۱۳۷۵، ۱۳۸۵ و آمارگیری جاری جمعیت ۱۳۷۰، آمار و اطلاعات اساسی مربوط به نیروی انسانی را جمع‌آوری نموده و علاوه بر آن از سال ۱۳۷۳، طرح نمونه‌گیری ویژگی‌های اشتغال و بیکاری خانوار را نیز به اجرا درآورده است. در سال‌های ۱۳۷۴ و ۱۳۷۵ به دلیل درگیری مرکز آمار ایران در اجرای سرشماری عمومی نفوس و مسکن این طرح اجرا نشد. نظر به نیاز برنامه ریزی کشور طرح مذکور مجدداً در پاییز سال‌های ۱۳۷۶ لغایت ۱۳۷۹ اجرا شد. در ادامه، طرح مذکور طی سال‌های ۱۳۸۰ لغایت ۱۳۸۲ در چهار فصل سال در ماه میانی هر فصل اجرا شد. در نهایت، در سال ۱۳۸۳ طرح مذکور در فصل بهار اجرا و نتایج آن منتشر شد. ولیکن در پاییز همان سال صرفاً برای دسترسی به شاخص‌های بازار کار کل کشور با نمونه‌ای محدود اجرا شد. از سال ۱۳۸۴ طرح آمارگیری از نیروی کار جایگزین طرح مذکور گردیده است. این طرح برای آمارگیری از نیروی کار ایران و بر اساس نیاز برنامه ریزان و با هدف شناخت ساختار و وضعیت جاری نیروی کار و تغییرات آن از طریق محاسبه‌ی شاخص‌های فصلی و سالانه‌ی نیروی کار، از جمله نرخ

مشارکت اقتصادی، نرخ بیکاری و برآورد تغییرات سالانه آن‌ها در سطح کل کشور و استان‌ها طراحی شده است. این آمارگیری منطبق بر آخرین توصیه‌های سازمان بین‌المللی کار^{۲۵} است و برای مقایسه‌ی بین‌المللی شاخص‌های نیروی کار، امکان بهتری را فراهم می‌کند و در حال حاضر مهمترین داده‌های مورد نیاز برای برآورد و استنباط آماری درباره‌ی پارمترهای نیروی کار ایران را فراهم می‌کند.

از نمونه‌گیری چرخشی برای بهبود برآورد تغییرات و افزایش دقت آمارها در این طرح استفاده می‌شود. برای دستیابی به برآوردهای فصلی و با توجه به ماهیت اقلام و سوال‌های پرسشنامه، آمارگیری در ماه میانی فصل در دو هفته‌ی مشخص انجام می‌شود و برخی سوالات پرسشنامه در مورد هفته‌ی قبل از آن دو هفته (هفته‌های مرجع) پرسیده می‌شود. همچنین در این طرح بدلیل ماهیت ادواری آن از نمونه‌ی پایه استفاده می‌شود. نمونه‌ی پایه، نمونه‌ای است که می‌توان از آن برای چند دوره از یک آمارگیری ادواری یا چند آمارگیری، زیر نمونه‌هایی انتخاب کرد. در ایران نمونه پایه این طرح برای سال‌های ۱۳۸۴ لغایت ۱۳۸۷ طی فرایندی از چارچوب سرشماری عمومی کارگاهی ۱۳۸۱ ساخته شده است که پس از ساخت نمونه‌ی پایه، فهرست‌برداری از خانوارهای مربوط به خوشه‌های نمونه پایه در پاییز ۱۳۸۳ انجام گرفت تا بر اساس آن بتوان خانوارهای نمونه را با توجه به الگوی چرخش موردنظر برای دوره‌های مختلف آمارگیری انتخاب نمود. نمونه پایه بعدی طرح از ابتدای سال ۱۳۸۸ براساس اطلاعات فایل سرشماری سال ۱۳۸۵ طراحی و مورد استفاده قرار گرفت. در نمونه پایه بعدی، از اطلاعات فایل سرشماری ۱۳۹۰ استفاده شده است.

از آنجا که هدف از انجام طرح نیروی کار بدست آوردن شاخص‌های اقتصادی مهمی مانند نرخ بیکاری، نرخ مشارکت اقتصادی و... است، لذا برای آشنایی بیشتر با این طرح در ادامه به معرفی برخی ویژگی‌های اصلی آن می‌پردازیم:

جامعه هدف^{۲۶}: مجموعه افرادی که طبق تعریف، عضو خانوارهای معمولی ساکن یا گروهی در نقاط شهری یا روستایی کشور هستند، جامعه هدف طرح را تشکیل می‌دهند. این طرح، اعضای خانوارهای معمولی غیر ساکن و خانوارهای موسسه ای را پوشش نمی‌دهد.

جامعه آمارگیری^{۲۷}: جامعه آمارگیری در این طرح به تعدادی از بلوک/آبادی‌ها در مناطق نظامی- انتظامی، مناطق صعب‌العبور و ناامن و تعدادی از آبادی‌های کمتر از پنج خانوار از جامعه هدف بدست آمده است.

واحد آماری^{۲۸}: واحد آماری، هر فرد عضو خانوارهای معمولی ساکن یا گروهی در نقاط شهری یا روستایی کشور است.

جمعیت فعال اقتصادی^{۲۹}: تمام افراد ۱۰ ساله و بیشتر (حداقل سن تعیین شده)، که در هفته تقویمی قبل از هفته آمارگیری (هفته مرجع) طبق تعریف کار، در تولید کالا و خدمات مشارکت داشته (شاغل) و یا از قابلیت مشارکت برخوردار بوده‌اند (بیکار)، جمعیت فعال اقتصادی محسوب می‌شوند.

²⁵ International Labour Organization (ILO)

²⁶ Target population

²⁷ Survey Population

²⁸ Statistical Unit

²⁹ Economically active population

جمعیت غیرفعال اقتصادی^{۳۰}: تمام افراد ۱۰ ساله و بیشتر که در طول هفته مرجع طبق تعریف، در هیچ یک از دو گروه شاغلان و بیکاران قرار نمی‌گیرند، جمعیت غیرفعال اقتصادی محسوب می‌شوند.
شاغل: تمام افراد ۱۰ ساله و بیشتر که در طول هفته مرجع، طبق تعریف کار، حداقل یک ساعت کار کرده و یا بنا به دلایلی بطور موقت کار را ترک کرده باشند، شاغل محسوب می‌شوند. شاغلان به طور عمده شامل دو گروه مزد و حقوق بگیران و خود اشتغالان می‌شوند.

بیکار^{۳۱}: بیکار به تمام افراد ۱۰ سال و بیشتر اطلاق می‌شود که:

۱. در هفته مرجع فاقد کار باشند.
 ۲. در هفته مرجع یا هفته بعد از آن آماده برای کار باشند.
 ۳. در هفته مرجع و سه هفته قبل از آن جویای کار باشند.
- افرادی که به دلیل آغاز به کار در آینده و یا انتظار بازگشت به شغل قبلی جویای کار نبوده، ولی فاقد کار و آماده برای کار بوده‌اند نیز بیکار محسوب می‌شوند.
دلسردان: به همه افرادی که شاغل نبوده و در هفته مرجع و سه هفته پیش از آن نیز به دلیل ناامیدی از پیدا کردن کار، جویای کار نبوده‌اند گفته دلسرد گفته می‌شود.

روش جمع‌آوری اطلاعات: روش آمارگیری در این طرح به صورت نمونه‌گیری است. در واقع در طرح آمارگیری نیروی کار، علاوه بر برآورد سطوح پارامترها (برآورد پارامترها در مقطع زمانی مشخص)، برآورد تغییرات پارامترها نیز موردنظر است. برای بدست آوردن هر دو نوع برآورد سطوح و تغییرات پارامترها با دقت مطلوب، از روش نمونه‌گیری چرخشی استفاده شده است. در این آمارگیری، اطلاعات از طریق مصاحبه رو در رو گردآوری می‌شود. به این ترتیب که مامور آمارگیر برای تکمیل پرسشنامه خانوار با مطلع‌ترین عضو خانوار و برای تکمیل پرسشنامه فردی با هریک از اعضای ده ساله و بیشتر خانوار مصاحبه می‌کند.

سطح برآورد^{۳۲}: سطح برآوردهای فصلی، "کل کشور" و سطح برآوردهای سالانه "کل کشور" و "استان" است.

پارامترهای مورد نظر: پارامترهای این طرح، شاخص‌های نیروی کار از جمله نرخ بیکاری، نرخ مشارکت اقتصادی و تعداد افراد بیکار و تغییرات آنها است. در زیر برخی از مهمترین این پارامترها معرفی شده‌اند:

✓ نرخ مشارکت اقتصادی (نرخ فعالیت)^{۳۳}: عبارتست از نسبت جمعیت فعال (شاغل و بیکار) به جمعیت در سن کار، ضربدر ۱۰۰.

³⁰ Economically inactive population

³¹ Unemployment

³² Estimated level

³³ Economic activity rate

$$\text{نرخ مشارکت اقتصادی} = \frac{\text{جمعیت فعال}}{\text{جمعیت در سن کار}} \times 100$$

✓ **نرخ اشتغال^{۳۴}**: عبارتست از نسبت جمعیت شاغل به جمعیت فعال (شاغل و بیکار)، ضربدر ۱۰۰.

$$\text{نرخ اشتغال} = \frac{\text{جمعیت شاغل}}{\text{جمعیت فعال}} \times 100$$

✓ **نسبت اشتغال**: عبارتست از نسبت جمعیت شاغل ۱۰ ساله و بیشتر (۱۵ ساله و بیشتر) به کل جمعیت ۱۰ ساله و بیشتر (۱۵ ساله و بیشتر) ضربدر ۱۰۰.

$$\text{نسبت اشتغال} = \frac{\text{جمعیت شاغل}}{\text{جمعیت در سن کار}} \times 100$$

✓ **نرخ اشتغال ویژه سنی**: عبارتست از نسبت جمعیت شاغل در یک گروه سنی معین به جمعیت فعال (شاغل و بیکار) همان گروه سنی ضربدر ۱۰۰.

$$\text{نرخ ویژه اشتغال سنی} = \frac{\text{جمعیت شاغل در گروه سنی معین}}{\text{جمعیت فعال گروه سنی}}$$

✓ **نرخ بیکاری^{۳۵}**: عبارتست از نسبت جمعیت بیکار به جمعیت فعال (شاغل و بیکار) ضربدر ۱۰۰.

$$\text{نرخ بیکاری} = \frac{\text{جمعیت بیکار}}{\text{جمعیت فعال}} \times 100$$

همان‌طور که در تعریف این شاخص می‌بینیم، در محاسبه نرخ بیکاری تنها دو مؤلفه جمعیت بیکار و جمعیت شاغل در نظر گرفته شده است، پس هر عاملی که باعث تغییر این دو مؤلفه شود، نرخ بیکاری را دستخوش تغییرات قرار می‌دهد. به‌عنوان مثال برای کاستن نرخ بیکاری دو راه وجود دارد:

۱. ایجاد فرصت‌های شغلی جدید: با ایجاد فرصت‌های جدید اشتغال در جامعه، تعدادی از افراد بیکار شاغل شده و در نتیجه از جمعیت بیکاران کاسته می‌شود و نهایتاً این کاهش، منجر به کاهش نرخ بیکاری می‌گردد.

³⁴ Employment Rate

³⁵ Unemployment Rate