

# به نام خدا

## روشهای نمونه گیری ۱

رشته: آمار

# فهرست مطالب

فصل اول: کلیات

فصل دوم: نمونه گیری تصادفی ساده

فصل سوم: نمونه گیری با طبقه بندی

فصل چهارم: نمونه گیری با احتمال متغیر



# فصل اول

## کلیات

## هدف های رفتاری فصل اول

- جامعه را تعریف کنید
- مفهوم سر شماری را بیان کنید
- نمونه را تعریف کنید
- بررسی های نمونه ای را رده بندی کنید
- هدف بررسی های توصیفی و بررسی های تحلیلی را بیان کنید
- مزایای نمونه گیری را توضیح دهید



- مراحل اصلی در یک نمونه گیری رایبان کنی
- انواع روش های برآورد را بنویسید -
- ملاک های یک برآورد خوب را توضیح دهید



- ملاک های نااریبی، سازگاری، کارایی  
و کفایت را برای یک برآورد تعریف کنید  
- برآورد فاصله ای را تعریف کنید  
- مواردی را که فرض نرمال بودن  
توزیع یک برآورد را می توان پذیرفت  
مشخص کنید

- ضریب تغییرات یک متغیر را محاسبه کنید

میانگین مربع خطا (**MSE**) را برای یک برآورد تعریف کرده و مقدار آن را تعیین نمایید

- مواردی که فرض نرمال بودن توزیع یک برآوردکننده را می توان پذیرفت مشخص کنید



-ضرب تغییرات یک متغیر را محاسبه کنید

-تفاوتهای نظریه ی نمونه گیری متعارف و

کلاسیک را بیان کنید



# مقدمه

در این فصل مفاهیم آماری لازم از  
قبیل جامعه، نمونه، نمونه گیری  
سرشماری، برآورد، فاصله اطمینان  
و... معرفی می شود



# ۱-۱ جامعه

در هر بررسی آماری، مجموعه عناصر مورد  
نظر را جامعه می گویند



# ۱-۲ سرشماری

سرشماری از جامعه متناهی بررسی  
است که تمام واحدهای جامعه را در بر  
گیرد



## ۱-۳ نمونه

نمونه بخشی از جامعه تحت بررسی است که با روشی که از پیش تعیین شده است انتخاب می شود، به قسمی که می توان از این بخش استنباط هایی درباره کل جامعه بدست آورد



## بررسی نمونه:

فرآیند انتخاب نمونه و استخراج نتایج و  
استنباطهای حاصل را بررسی نمونه ای می نامند



۱- بررسی نمونه توصیفی: هدف صرفاً کسب اطلاع در مورد گروه های بزرگ است

۲- بررسی نمونه تحلیلی: بین زیرگروههای متفاوتی از جامعه، برا کشف تاوتهای انها مقایسه هایی صورت می گیرد و یا فرض هایی را درباره ی دلائل این تفاوتها عنوان کرده و موردتحقیق قرار می دهند



# ۱-۴ مزایای نمونه گیری

۱- **تقلیل هزینه:** اگر داده ها فقط از نسبت کوچکی از توده ی جامعه تامین شوند مسلماً هزینه ی تهیه ی آنها به مراتب کمتر از سرشماری است.

۲- **سرعت بیشتر:** چون حجم نمونه کمتر از جامعه در سرشماری است، جمع آوری و تلخیص داده ها با سرعت بیشتر، یعنی با وقت کمتری انجام می شود.



۳- قدرت عمل بیشتر

۴- صحت عمل بیشتر

۵- حفظ واحدهای جامعه: در بعضی از جوامع امکان سرشماری نیست و ناگزیریم برای بررسی مشخصه‌ی مورد نظر از نمونه‌گیری استفاده کنیم.





# ۱-۵ مراحل اصلی در یک بررسی نمونه ای

## ۱- اهداف بررسی

همواره باید حکمی روشن و صریح درباره هدف های بررسی در دست باشد

## ۲- جامعه نمونه گیری

جامعه ای که از آن نمونه می گیریم باید دقیق تعریف شود. جامعه ای که می خواهیم درباره ی آن کسب اطلاع کنیم جامعه هدف نامیده می شود. جامعه ای که از آن نمونه گیری می کنیم قاعدتاً باید بر جامعه هدف منطبق باشد ولی بی دلایل عملی و یا برای سهولت کار جامعه مورد نمونه گیری اغلب محدود تر از جامعه هدف است

۳- جمع آوری داده ها:

لازم است تحقیق کنیم که تمام داده ها به اهداف بررسی مربوط اند و هیچ داده اساسی از قلم نیفتاده است

۴ - درجه دقت مطلوب

۵ - روش های اندازه گیری



۶- چارچوب

فهرست واحد های نمونه را چارچوب می نامند که غالباً تعیین آن یکی از مسائل عمده ی کار نمونه گیری است

۷- انتخاب نمونه

۸- پیش آزمون

پیش آزمون معمولاً با هزینه کم مانع از به هدر رفتن هزینه زیاد نمونه گیری اصلی می شود

۹- آموزش آمارگران

۱۰- تلخیص و تحلیل داده ها

۱۱- اطلاعات حاصل برای بررسی های آتی  
هر نمونه ای که از جامعه گرفته می شود، بالقوه  
راهنمایی برای اصلاح نمونه گیری های بعدی  
است



۱-۶ مروری بر برخی مفاهیم آماری  
در نظریه نمونه گیری، برآورد پارامترها و تهیه ی  
حکم احتمالاتی درباره ی دقت برآورد از اهداف  
اصلی است.



# ۱- برآورد

برآوردکننده و یا برآوردی را از نظر آماری  
خوب گوئیم که در چهار ملاک  
نااریبی، سازگاری، کارایی و کفایت صادق باشد



# نا اریبی

اگر پارامتر نا معلوم جامعه  $\theta$  باشد و اگر  
برآورد کننده ی آن  $\hat{\theta}$ ، آماره ای باشد که  
میانگین آن  $\theta$  شود در اینصورت  $\hat{\theta}$   
برآورد کننده ی نااریب  $\theta$  گویند و داریم

$$E(\hat{\theta}) = \theta$$

$$E(\hat{\theta}) = \theta + \alpha \quad \text{واگر}$$

در اینصورت  $\hat{\theta}$  را برآورد اریب  $\theta$  گوئیم  
 $\alpha$  را اریبی برآورد کننده گویند



# سازگاری

به ازای  $\varepsilon$  بی نهایت کوچک مثبت

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < \varepsilon) \rightarrow 1$$

جامعه را نرمال در نظر بگیریم، چون در جامعه نرمال  
نگین و میانه یکی است میانه نمونه هم برآوردکننده  
سازگاری برای  $\bar{Y}_N$  خواهد بود





# کارایی

اگر برای پارامتر  $\theta$  ی جامعه، دو برآورد کننده ی  $\hat{\theta}_1$  ،  $\hat{\theta}_2$  داشته باشیم و  $\text{var}(\hat{\theta}_1) \neq \text{var}(\hat{\theta}_2)$  آنگاه کارایی  $\hat{\theta}_2$  نسبت به  $\hat{\theta}_1$  بصورت زیر است:

$$e = \frac{\text{var}(\hat{\theta}_1)}{\text{var}(\hat{\theta}_2)} \quad 0 \leq e \leq 1$$



## 2- برآورد فاصله ای، فاصله اطمینان

در بررسی های نمونه ای، روش اصلی برآورد غالباً مبتنی بر برآورد فاصله ای است. یک برآورد فاصله ای برای پارامتر  $\theta$  ی جامعه، فاصله ای بصورت  $K_1 < \theta < K_2$  است که در آن  $K_1$  و  $K_2$  به توزیع  $\hat{\theta}$  و به مقداری که برآوردکننده ی  $\hat{\theta}$  در نمونه ای مفروض اختیار میکند بستگی دارد. می توانیم از توزیع  $\hat{\theta}$  استفاده کنیم و  $K_1$  و  $K_2$  را طوری انتخاب کنیم که به ازای هر احتمال مشخص  $1 - \alpha$  احتمال همراه با فاصله ی  $K_1 < \theta < K_2$  برابر  $1 - \alpha$  باشد. این فاصله را فاصله اطمینان  $100(1 - \alpha)\%$  می نامند و  $1 - \alpha$  را ضریب اطمینان می گویند.  $K_1$  و  $K_2$  حدود اطمینان اند.

## ۳- میانگین مربع خطا

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + \beta^2$$

مربع اریبی + واریانس

اگر  $MSE$  را به عنوان ملاک  
دقت یک برآورد کننده به کار

بریم به این نتیجه می رسیم

که دو برآورد کننده ای که


دارای  $MSE$  برابر هستند هم

ارزند.



## ۴-اعتبار تقریب نرمال

ثابت شده است که برای هر جامعه با انحراف معیار متناهی، توزیع میانگین نمونه وقتی  $n$  زیاد می شود، به توزیع نرمال می گراید  
برای انتخاب حجم نمونه از  $n > 2\sigma G_1^2$  استفاده می کنیم:


$$G_1 = \frac{E(y_i - \overline{y_N})^3}{\sigma^3} = \frac{1}{N\sigma^3} \sum_{i=1}^N (y_i - \overline{y_N})^3$$

# ۱-۷ ضریب تغییرات

$$C^2 = \frac{1}{N} \sum_{i=1}^N \left[ \frac{y_i - \bar{y}_N}{\bar{y}_N} \right]^2 = \frac{1}{N \bar{y}_N^2} \sum_{i=1}^N (y_i - \bar{y}_N)^2 = \frac{\sigma^2}{\bar{y}_N^2}$$

$$C = \frac{\sigma}{\bar{y}_N}$$



ضریب تغییرات



واریانس نسبی جامعه یا واریانس نسبی توزیع  $y$   
متغیر تصادفی

ضریب تغییرات، مقایسه ی تغییرات میانگین های دو  
جامعه متفاوت را با هم مقایسه می سازد



ویژگی مهم C  
 در بسیاری از پدیده ها مقدارش تقریباً ثابت  
 است

$$\frac{C(y)}{\sqrt{n}} = \frac{\sigma_y}{y_N \sqrt{n}} = \frac{\sqrt{Var(\bar{y}_n)}}{E(\bar{y}_n)} = C(\bar{y}_n)$$

ضریب تغییر میانگین نمونه ای به حجم  
 میانگین جامعه است.

$$\frac{1}{\sqrt{n}}$$

برابر



# ۱-۸ تفاوت‌های نظریه نمونه گیری

## متعارف و کلاسیک

یکی آنکه، در نظریه ی کلاسیک، فرض بر این است که اندازه های واحد ها نمونه از توزیع فراوانی معینی پیروی می کنند ولی در نظریه ی نمونه گیری معمولاً درباره ی توزیع فراوانی اطلاعات چندانی در دست نیست

دوم اینکه برای نمونه گیری کلاسیک حجم جامعه غالباً نامتناهی است ولی در نظریه نمونه گیری مورد بحث کتاب، حجم جامعه متناهی است



# فصل دوم

## نمونه گیری تصادفی ساده



# هدف های رفتاری فصل دوم

-نمونه گیری تصادفی ساده را تعریف کنید.

-نمونه گیری تصادفی ساده با جایگذاری را تعریف کنید.

-نمونه گیری تصادفی ساده بدون جایگذاری را تعریف کنید.

-جدول اعداد تصادفی را برای انتخاب نمونه تصادفی ساده به کار ببرید.

-ویژگی های مهم نمونه گیری تصادفی بدون جایگذاری را بیان کنید.

-مشخصه ی جامعه را تعریف کنید.



- ثابت کنید که میانگین نمونه ی تصادفی ساده بدون جایگذاری برآورد کننده ی ناریب میانگین جامعه است.
- ثابت کنید که  $s^2$ ، واریانس نمونه تصادفی ساده بدون جایگذاری برآورد کننده ی ناریب  $s^2$  ی جامعه است.
- برآورد ناریب واریانس میانگین نمونه را در یک نمونه گیری تصادفی ساده بدون جایگذاری بدست آورید.
- ثابت کنید  $N \overline{y_n}$  برآورد کننده ناریب  $t_N$ ، مجموع واحدهای جامعه است.
- برآورد کننده ناریب واریانس  $\hat{t}_N$  را بدست آورید.
- ضریب تصحیح واریانس میانگین نمونه را برای جامعه متناهی تعریف و مواردی که این ضریب را نادیده می گیرند بیان کنید.



- موارد استفاده از توزیع  $t$  یا توزیع نرمال رادر برآورد فاصله ای میانگین جامعه تشخیص دهید.

- نمونه گیری تصادفی ساده با جایگذاری را تعریف کنید.

- ثابت کنید که در نمونه گیری تصادفی با جایگذاری،  $s^2$  ی نمونه برآورد کننده ی ناریب  $\sigma^2$  ی جامعه است.

- نسبت  $P$  را در جامعه با استفاده از نسبت نمونه ای  $p$  بیان کنید.

- واریانس نسبت جامعه را بدست آورید.

- برآورد ناریب واریانس  $P$  رابدست آورید.





- حدود اطمینان برای نسبت جامعه،  $P$ ، را محاسبه کنید.
- مواردی را که ممکن است از توزیع نرمال به عنوان تقریبی خوب برای محاسبه حدود اطمینان  $P$  استفاده کرد مشخص کنید.
- نسبت یک رده را در جامعه ای با چند رده محاسبه کنید.



# مقدمه

یکی از روش های متداول در نمونه گیری که برای جمع آوری اطلاعات به کار میرود، نمونه گیری تصادفی ساده است که بصورت "باجایگذاری" و "بدون جایگذاری" انجام می شود



## ۲-۱ نمونه گیری تصادفی ساده

روش انتخاب  $\underline{n}$  واحد مشخص از جامعه ای به حجم  $\underline{N}$  واحد است، به

قسمی که همه ی  $\binom{N}{\underline{n}}$  نمونه ای که می توان انتخاب کرد شانس یکسانی برای انتخاب شدن داشته باشد.



\*احتمال اینکه  $n$  واحد مشخص در  
استخراج انتخاب شوند برابر  $n$  است با

$$\frac{1}{\binom{N}{n}}$$

چون هر عددی که در استخراج های متوالی  
بدست می آید کنار گذاشته می شوند، روش  
مزبور را در نمونه گیری تصادفی **بدون**

**جایگذاری** می نامند



\*نمونه گیری تصادفی با جایگذاری هم  
قابل انجام است که در هر استخراج شانس  
در آمدن هر یک از واحد های جامعه  
مقدار  $\frac{n}{N}$  است

دقت نمونه گیری با جایگذاری از بدون جایگذاری کمتر است





## ۲-۲ انتخاب نمونه تصادفی ساده

جداول اعداد تصادفی جداولی از ارقام

۰، ۱، ۲، ۳، ۴، ۵، ۶، ۷، ۸، ۹

هستند که شانس انتخاب تمام ارقام در هر استخراج یکی است. از این جداول برای انتخاب نمونه تصادفی ساده استفاده میشود.

## ۲-۳ ویژگی مهم نمونه گیری تصادفی بدون جایگذاری

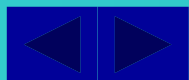
احتمال استخراج یک واحد مشخص در انتخاب  $r$  ام

برابر  $\frac{1}{N}$  است.



## ۲-۴ تعاریف و نما دها

در یک بررسی نمونه ای، توجه ما به برخی ویژگی های واحدهاست.  
هر ویژگی موردنظر را یک **مشخصه** یا یک **صفت می** نامیم.



—  
 $\bar{y}_N$  برآورد کننده  $\bar{y}_N$  → عدد ثابت → میانگین جامعه  $y_N$   
—

$y_n$  → متغیر تصادفی → میانگین نمونه  
برای نمونه ای مشخص با  $y_n$  مقداری است ثابت

$t_N$  → عدد ثابت → مجموعه واحدهای جامعه  $t_N$   
 $\hat{t}_N =$  برآورد کننده  $t_N$   
 $t_n$  → متغیر تصادفی → مجموعه واحدهای نمونه  $t_n$   
برای نمونه ای مشخص با

$t_n$  مقداری است ثابت

$$\bar{y}_N = \frac{1}{N} \sum_{i=1}^N y_i \quad \bar{y}_n = \frac{1}{n} \sum_{i=1}^n y_i$$



## ۲-۵ قضیه

$\bar{y}_n$ ، میانگین نمونه تصادفی ساده بدون جایگذاری، برآورد کننده ی نااریب  $\bar{y}_N$ ، میانگین جامعه ست،

یعنی

$$E(\bar{y}_n) = \bar{y}_N$$



## ۲-۶ واریانس میانگین نمونه

قبل از اینکه واریانس میانگین نمونه تصادفی ساده را با بیان قضیه مشخص کنیم خاطر نشان می سازیم که واریانس جامعه  $\underline{N}$  واحدی را با  $\sigma^2$  و تغییرات این جامعه را با  $S^2$  نشان می دهیم.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y}_N)^2$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y}_N)^2$$



برای محاسبه واریانس  $\bar{y}_n$  به لم زیر نیاز داریم

از جامعه ای به حجم  $N$ ، نمونه تصادفی ساده بدون جایگذاری به حجم  $n$  را انتخاب می کنیم. اگر  $y_i$  و  $y_j$  دو واحد مشخص نمونه در دو انتخاب متوالی نمونه باشند و  $\sigma^2$  واریانس جامعه فرض شود، آنگاه

$$Cov(y_i, y_j) = \frac{-\sigma^2}{N-1}$$



## ۲-۷ قضیه

اگر  $y_n$  میانگین نمونه تصادفی ساده بدون جایگذاری از جامعه به حجم  $N$  باشد آنگاه

$$Var(\bar{y}_n) = \left[ \frac{1}{n} - \frac{1}{N} \right] S^2 = (1 - f) \frac{S^2}{n}$$

که در آن  $S^2$  تغییرات جامعه و  $f = \frac{n}{N}$  کسر نمونه گیری است.



## ۲-۸ قضیه

اگر  $S^2$  واریانس یک نمونه تصادفی ساده  
ی بدون جایگذاری به حجم  $n$  از  
جامعه ای به حجم  $N$  باشد، آنگاه  
بر آورد کننده ی نااریب  $S^2$  جامعه  
است. یعنی

$$E(s^2) = S^2$$





## فرع ۱.

بر آورد کننده نااریب  $Var(\bar{y}_n)$  در یک  
نمونه گیری تصادفی ساده بدون جایگذاری  
بصورت

$$\hat{Var}(\bar{y}_n) = \left[ \frac{1}{n} - \frac{1}{N} \right] s^2 = (1-f)s^2$$

است که در آن،  $S^2$  واریانس نمونه است

## فرع ۲.

می توان نتیجه گرفت که

$$\sigma_{\bar{y}_n} = \sqrt{\frac{1}{n} - \frac{1}{N}} s = \sqrt{\frac{1-f}{n}} \cdot s$$



## فرع ۳.

### بنابر رابطه فرع ۱

$$\hat{Var}(\bar{y}_n) = \left[ \frac{1}{n} - \frac{1}{N} \right] s^2 = (1-f) \frac{s^2}{n}$$

برآورد کننده نااریب  $Var(\bar{y}_n)$  است، اما اگر از طرفین این برابری ها جذر بگیریم نتیجه می شود

$$\hat{\sigma}(\bar{y}_n) = \sqrt{\frac{1}{n} - \frac{1}{N}} \cdot s = \sqrt{\frac{1-f}{n}} \cdot s$$

وقتی جامعه مورد نمونه گیری توزیع نرمال داشته باشد می توان برآورد کننده ای نا اریب برای  $\bar{y}_n$  یافت.

فرع ۴.

$N\bar{y}_n$  برآورد کننده ی نا اریب  $t_N$  است.

$$E(N\bar{y}_n) = t_N \Rightarrow \hat{t}_N = N\bar{y}_n$$



## فرع ۵ .

واریانس  $\hat{t}_N = N\bar{y}_n$  یعنی واریانس  
برآورد کننده مجموع واحد های  
جامعه برابر است با

$$Var(\hat{t}_N) = \left[ \frac{1}{n} - \frac{1}{N} \right] N^2 S^2 = (1 - f) \frac{N^2 S^2}{n}$$



## ۲-۹ ضرب تصحیح برای جامعه متناهی

برای نمونه تصادفی به حجم  $n$  از جامعه نا متناهی

$$\text{داریم} \quad \text{Var} \left( \bar{y}_n \right) = \frac{\sigma^2}{n}$$

که در آن  $\sigma^2$ ، واریانس جامعه است. وقتی جامعه متناهی است دیدیم که بر حسب  $\sigma^2$

$$\text{Var} \left( \bar{y}_n \right) = \frac{\sigma^2}{n} \cdot \frac{N - n}{N - 1}$$



## ۲-۱۰ حدود اطمینان میانگین جامعه

معمولاً فرض می کنند که  $\bar{y}_n$  حول  $\bar{y}_N$  و  $t_n$  حول  $t_N$  توزیع نرمال دارند

بنا براین داریم

$$\begin{cases} \bar{y}_L = \bar{y}_n - \frac{zNS}{\sqrt{n}} \sqrt{1-f} \\ \bar{y}_U = \bar{y}_n + \frac{zNS}{\sqrt{n}} \sqrt{1-f} \end{cases}$$

حد بالا برای میانگین جامعه =  $\bar{y}_L$   
حد پایین برای میانگین جامعه =  $\bar{y}_U$

$$\begin{cases} t_L = \bar{y}_n - \frac{zS}{\sqrt{n}} \sqrt{1-f} \\ t_U = \bar{y}_n + \frac{zS}{\sqrt{n}} \sqrt{1-f} \end{cases}$$

حد پایین برای مجموع واحدهای جامعه =  $t_L$   
حد بالا برای مجموع واحدهای جامعه =  $t_U$



# ۲-۱۱ نمونه گیری تصادفی ساده با جایگذاری

اگر در انتخاب  $\underline{n}$  واحد نمونه ، پس از انتخاب هر واحد آن را به جامعه برگردانیم و انتخاب بعدی را انجام دهیم نمونه گیری تصادفی ساده را با جایگذاری می نامند.

$$\frac{1}{N^n} = \underline{n} \text{ واحدی}$$



## ۲-۱۲ میانگین نمونه تصادفی ساده با جایگذاری

فرض می کنیم واحد  $y_i$  ی جامعه ،  
 $Z_i$  بار ظاهر شود. بدیهی است  $Z_i$  می  
تواند از ۰ تا  $n$  تغییر کند. با این فرض

$$\bar{y}_n = \frac{1}{n} \sum_{i=1}^N Z_i y_i$$





## ۲-۱۳ قضیه

میانگین نمونه ی تصادفی ساده با جایگذاری  
برآورد کننده ی نااریب میانگین جامعه است.

یعنی،

$$E(\bar{y}_n) = \bar{y}_N$$



## ۲-۱۴ واریانس میانگین نمونه ی تصادفی ساده با جایگذاری

تعداد میانگین های ممکن برابر  $N^n$  است. اگر پراکندگی این میانگین ها کم باشد، میانگین که در عمل از روی یک نمونه حاصل می شود با  $\bar{y}_N$  خیلی فاصله ندارد و لذا برآورد نا اریب مطلوبی است که برای تشخیص زیاد و کم بودن این پراکندگی  $Var(\bar{y}_n)$  را محاسبه می کنیم

$$Var(Z_i) = n \cdot \frac{1}{N} \left( 1 - \frac{1}{N} \right) \quad Cov(Z_i, Z_j) = \frac{-n}{N^2}$$

$$Var(\bar{y}_n) = \frac{\sigma^2}{n} \Rightarrow \frac{\sigma}{\sqrt{n}} \quad \text{انحراف معیار } \bar{y}_n \text{ برابر}$$

## ۲-۱۵ قضیه

در نمونه گیری تصادفی ساده با جایگذاری  
 $S^2$  نمونه، برآورد کننده ی نا اریب  $\sigma^2$

ی جامعه است یعنی  $E(s^2) = \sigma^2$

**نتیجه:**

چون  $S^2$  برآورد کننده ی نا اریب  $\sigma^2$  است پس  $Var(\bar{y}_n) = \frac{s^2}{n}$

لذا پراکندگی  $\bar{y}_n$  های ممکن را با  $\frac{s^2}{n}$  حاصل از یک  
نمونه برآورد می کنیم.



## ۲-۱۷ نماد گذاری

دو رده را با  $C$  و  $\bar{C}$  نشان می دهیم.  
هر واحد جامعه را مطابق معمول  $N$  می گیریم.  
به روش تصادفی ساده، بدون جایگذاری، نمونه ای به  
حجم  $n$  تهیه می کنیم. نماد های زیر را اختیار می کنیم:

$A$	تعداد واحد های جامعه که در رده ی
$a$	تعداد واحد های نمونه که در رده
$P = \frac{A}{N}$	نسبت واحد هایی از جامعه که در رده $C$
$p = \frac{a}{n}$	نسبت واحدهایی از نمونه که در رده $C$

## ۲-۲۲ قضیه

برآورد کننده ی نااریب  $P$ ، برابر  $p$  است.

نتیجه ۱:

اگر نسبت واحدهایی از جامعه را که در رده ی  $\bar{C}$  اند

$Q$  نشان می دهیم  $P+Q=1$

و اگر نسبت واحدهایی از نمونه را که در رده ی  $\bar{C}$

هستند با  $q$  نشان می دهیم  $p+q=1$



نتیجه ۲:

بدیهی است که  $q$  برآوردکننده ی

نااریب  $Q$  است.

$$\hat{A} = a \cdot \frac{N}{n}$$



## ۲-۱۹ محاسبه تغییرات جامعه و نمونه

$$\sum_{i=1}^N y_i^2 = \text{مجموع مربعات یکها و صفرها در دنباله جامعه}$$

$$A = NP = \text{مجموع مربعات یکها در دنباله جامعه}$$

$$\sum_{i=1}^n y_i^2 = \text{مجموع مربعات یکها و صفرها در دنباله جامعه}$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y}_N)^2 = \frac{NPQ}{N-1} \quad a = np$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_n)^2 = \frac{n}{n-1} pq$$

۲-۲۰ قضیه

واریانس  $P$  بصورت زیر است.

$$Var(p) = \frac{N-n}{N-1} \cdot \frac{PQ}{n}$$

۲-۲۱ برآورد واریانس

در عمل برای تعیین دقت برآورد کننده  $Q$ ،  $P$



و  $p$  را نداریم لذا باید برآوردی برای

$Var(p)$

فاهم کنیم.



## ۲-۲۲ قضیه

برآورد کننده ی نا اریب  $Var(p)$  بصورت زیر است

$$\hat{Var}(\bar{y}_n) = \hat{Var}(\hat{P}) = \frac{N - n}{n - 1} \frac{Pq}{N}$$



## نتیجه ۱

اگر  $N$  خیلی بزرگ باشد و به قسمی که از  $fpc$  بتوان صرف نظر کرد

$$\hat{Var}(p) = \frac{pq}{n-1}$$

## نتیجه ۲

بر آورد نااریب واریانس  $\hat{A} = Np \leftarrow \hat{Var}(\hat{A}) = \frac{N-n}{n-1} \cdot Npq$

$$\hat{Var}(\hat{A}) \cong \frac{N^2 pq}{n-1}$$

اگر حجم جامعه بزرگ



## ۲-۲۳ حدود اطمینان $P$

حدود تقریبی  $P$  رادر سطح اطمینان  $1 - \alpha$  بصورت زیر بدست می آوریم

$$\left\{ \begin{array}{l} \text{حد پایین} = p - z \sqrt{\frac{N - n}{n - 1} \frac{pq}{N}} \\ \text{حد بالا} = p + z \sqrt{\frac{N - n}{n - 1} \frac{pq}{N}} \end{array} \right.$$



## ۲-۲۴ تعیین نسبت در جامعه ای با چند رده

اگر  $N$  واحد جامعه به جای دو رده در  $K$  رده قرار گیرند تعداد واحدهای رده  $i$  را  $A_i$  و

نسبت این واحدها را در جامعه برابر  $P_i$  می گیریم

$$p_i = \frac{a_i}{n}$$



## ۲-۲۵ برآورد حجم نمونه

در طرح ریزی بررسی نمونه ای، اخذ تصمیم درباره حجم نمونه از نظرتامین دقت نتایج نمونه گیری و صرفه جویی در میزان وقت و هزینه ی آن از اهمیت خاصی برخوردار است. بدیهی است بزرگ بودن حجم نمونه موجب صرف هزینه و وقت زیاد، و کوچک بودن آن موجب عدم دقت کافی برآورد ها است



## ۲-۲ محاسبه حجم نمونه برای داده های پیوسته

اگر ضریب تغییرات جامعه  $\approx \frac{\overline{y}_n}{\overline{y}}$

$$n = \left[ \frac{tS}{r\overline{y}_N} \right]^2 / \left[ 1 + \frac{1}{N} \left[ \frac{tS}{r\overline{y}_N} \right]^2 \right] \quad \text{آنگاه}$$

$$\gamma = \left( \frac{r}{t} \right)^2 \quad \text{که} \quad n_o = \left[ \frac{tS}{r\overline{y}_N} \right]^2 = \frac{1}{\gamma} \left[ \frac{S}{\overline{y}_N} \right]^2 \quad \text{اگر } N \text{ بزرگ باشد}$$

$$n = \frac{n_o}{1 + n_o / N}$$

که با جایگذاری حجم نمونه برابر است

**نتیجه**



## برآورد حجم نمونه

$$\hat{n} = \left[ \frac{t\hat{S}}{r\hat{\bar{y}}_N} \right]^2 / \left[ 1 + \frac{1}{N} \left[ \frac{t\hat{S}}{r\hat{\bar{y}}_N} \right]^2 \right]$$



۲-۲۷ محاسبه حجم نمونه بر حسب فاصله اطمینان مفروض

گاهی حجم نمونه را به قسمی تعیین می کنند که فاصله اطمینان میانگین جامعه با ضریب اطمینان معلوم  $1 - \alpha$  به طول مفروض  $2l$  باشد

$$\hat{n} = (1 - f) \frac{s^2}{l^2} t^2$$





۲-۲۸ تعیین حجم نمونه در نمونه گیری برای نسبت ها

$$n = \frac{t^2 PQ / d^2}{1 + \frac{1}{N} \left[ \frac{t^2 PQ}{d^2} - 1 \right]}$$

بر آورد  $n$   $\hat{n} = \frac{t^2 pq / d^2}{1 + \frac{1}{N} \left[ \frac{t^2 pq}{d^2} - 1 \right]}$

$$\hat{n} = \frac{pq/v}{1 + \frac{1}{N} \left[ \frac{pq}{v} - 1 \right]}$$

اگر قرار دهیم  $\frac{d^2}{t^2} = v$  ←



وقتی  $N$  بزرگ است  $\hat{n}_o = \frac{pq}{v}$

$$n = \frac{\hat{n}_o}{1 + \frac{\hat{n}_o - 1}{N}} \approx \frac{\hat{n}_o}{1 + \frac{\hat{n}_o}{N}}$$



۳۰-۲ قضیه

اگر  $(x, y)$  یک زوج متغیر تصادفی باشند و نمونه تصادفی ساده ای به حجم  $n$  از جامعه  $(x, y)$  که به حجم  $N$  است انتخاب کنیم و اگر  $\bar{x}_n$  و  $\bar{y}_n$  میانگین های نمونه ای  $x$  و  $y$  باشند

$$Cov(\bar{x}_n, \bar{y}_n) = E[(\bar{y}_n - \bar{y}_N)(\bar{x}_n - \bar{x}_N)] = \left[ \frac{1}{n} - \frac{1}{N} \right] S_{xy} \quad \text{آنگاه}$$

که در آن  $\bar{x}_N$  و  $\bar{y}_N$  به ترتیب میانگین های جامعه ای  $x$  و  $y$  باشند

$$S_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_N)(y_i - \bar{y}_N) \quad \text{و}$$



## ۲-۳۱ قضیه

اگر ضریب همبستگی بین زوج  $(x, y)$  برابر  $\rho$

باشد و اگر  $\bar{y}_n$  و  $\bar{x}_n$  میانگین های نمونه ای به

حجم  $n$  از زوج  $(x, y)$  باشند آنگاه ضریب

همبستگی زوج  $(\bar{x}_n, \bar{y}_n)$  به حجم نمونه بستگی

ندارد و برابر با  $\rho$  است.



# فصل سوم

## نمونه گیری با طبقه بندی

## هدف های رفتاری فصل سوم

- نمونه گیری تصادفی با طبقه بندی را تعریف کنید
- سه دلیل عمده طبقه بندی جامعه را بیان کنید
- نمونه گیری با تخصیص متناسب را تعریف کنید
- ثابت کنید که  $\bar{Y}_{st}$  برآورد نااریب میانگین جامعه است
- واریانس  $\bar{Y}_{st}$  را در صورتی که نمونه هاتی طبقات مستقل از هم باشند بدست آورید



- واریانس  $\bar{Y}_{st}$  را در حالت تخصیص متناسب بدست آورید  
- ثابت کنید که برآورد ناریب مجموع واحدهای جامعه  $N\bar{Y}_{st}$  است.

- برآورد کننده واریانس  $\bar{Y}_{st}$  را در نمونه گیری تصادفی ساده با طبقه بندی بدست آورید  
- حدود اطمینان میانگین و مجموع واحدهای جامعه را در نمونه گیری تصادفی ساده با طبقه بندی بدست آورید



- حدود اطمینان میانگین و مجموع واحدهای جامعه را در نمونه گیری تصادفی ساده با طبقه بندی بدست آورید
- حجم نمونه را در طبقات مختلف طوری بدست آورید که هزینه نمونه گیری مینیمم شود
- حجم نمونه را در طبقات مختلف طوری بدست آورید که واریانس  $\bar{Y}_{st}$  مینیمم شود
- حجم نمونه را در طبقات مختلف، برای حالتی که هزینه ی نمونه گیری برای هر واحد در تمام طبقات یکسان است بدست آورید

.....





## مقدمه

نمونه گیری تصادفی با طبقه بندی نیز یکی از روش های متداول نمونه گیری است که در این فصل با آن آشنا خواهید شد. با اطلاعات حاصل از این روش ، برآوردهایی برای مشخصه های جامعه و حجم نمونه لازم در هر طبقه بدست می آید.



## ۳-۱ تعریف.

در نمونه گیری با طبقه بندی، جامعه به حجم  $N$  را ابتدا به زیر جامعه هایی به حجم های  $N_1, N_2, \dots, N_L$  تقسیم می کنند. این زیر جامعه ها متداخل نیستند و اجتماع آنها برابر با کل جامعه است، یعنی

$$N_1 + N_2 + \dots + N_L = N$$

هر زیر جامعه را یک طبقه می نامند

- حجم های نمونه های طبقات را با  $n_1, n_2, \dots, n_L$  نشان می دهیم
- اگر از هر طبقه نمونه ای به روش تصادفی ساده گرفته شود شیوه ی کلی نمونه گیری را، نمونه گیری تصادفی با طبقه بندی می نامند.



## ۳-۲ نمادها و تعاریف

زیر نویس  $\underline{h}$  معرف طبقه است که از  $\underline{1}$  تا  $\underline{L}$  تغییر می کند.  $\underline{L}$  تعداد طبقات است.

زیر نویس  $\underline{i}$  برای تعیین شماره ی واحد ها در داخل هر طبقه به کار می رود.  $\underline{N}$  تعداد کل افراد جامعه است. نمادهای زیر برای طبقه  $\underline{h}$  و  $\underline{h}=1,2,\dots,L$  تعریف می شود.



## ۲-۳ نمادها و تعاریف

تعداد کل واحدها در طبقه  $h$  ام

$$N_h$$

$$n_h$$

تعداد واحد های نمونه ی طبقه ی  $h$  ام

برای  $n_h$   $i = 1, 2, \dots, n_h$  مقدار واحد  $i$  ام

$$Y_{hi}$$

در طبقه ی  $h$  ام

$$f_h = \frac{n_h}{N_h}$$

$h$  ام

کسر نمونه گیری برای طبقه

وزن طبقه ی  $h$  ام

$$W_h = \frac{N_h}{N}$$





## ۲-۳ نمادها و تعاریف

$$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} Y_{hi}$$

میانگین طبقه  $h$  ام

$$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} Y_{hi}$$

میانگین نمونه طبقه  $h$  ام

$$S_h^2 = \frac{1}{(N_h - 1)} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$$

تغییرات طبقه  $h$  ام

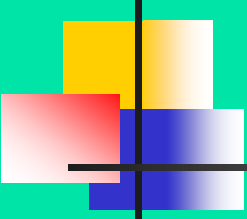
$$s_h^2 = \frac{1}{(n_h - 1)} \sum_{i=1}^{n_h} (Y_{hi} - \bar{y}_h)^2$$

تغییرات نمونه ای طبقه  $h$  ام




میانگین های موزون میانگین های نمونه ای طبقات

را به نام میانگین با طبقه بندی می نامیم


$$\bar{Y}_{st} = \frac{1}{N} \sum_{h=1}^L N_h \bar{y}_h = \sum_{h=1}^L W_h \bar{y}_h$$

این میانگین با میانگین نمونه ی  $n$  واحدی حاصل از روی به هم ریختن نمونه های طبقات، در حالت کلی، تفاوت دارد. بدیهی است که وقتی  $\bar{y}_{st}$  با میانگین نمونه کلی یکی است که در هر طبقه  $\frac{n_h}{N} = \frac{n}{N}$  یا  $f_h = f$  باشد  $(h = 1, 2, 3, \dots, L)$



یعنی کسر نمونه گیری در تمام طبقات یکسان باشد. نمونه گیری با طبقه بندی را وقتی با این ویژگی همراه باشد می گویند با تخصیص متناسب است

## ۳-۳ قضیه

اگر در هر طبقه،  $\bar{y}_h$  برآورد کننده نااریب میانگین طبقه  $h$  باشد، آنگاه  $\bar{y}_{st}$  برآورد کننده نااریب میانگین کل جامعه یعنی برآورد کننده نااریب  $\bar{y}_N$  است.

$$E(\bar{y}_{st}) = \bar{y}_N$$

پس  $\hat{\bar{Y}}_N = Y_{st}$  که این برآورد کننده ویژگی نااریبی را هم دارد



### ۳-۴ قضیه

اگر نمونه های طبقات، مستقل از هم گرفته شوند

$$Var(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 Var(\bar{y}_h)$$

که در آن  $Var(\bar{y}_h)$ ، واریانس  $\bar{y}_h$  برای تمام نمونه های ممکن از طبقه  $h$  ام است.

### ۳-۵ قضیه

برای نمونه گیری تصادفی با طبقه بندی واریانس  $\bar{Y}_{st}$  بصورت زیر است.

$$Var(\bar{Y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h} = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h} - \sum_{h=1}^L \frac{W_h S_h^2}{N}$$



## فرع ۱

اگر در تمام طبقات کسر نمونه گیری، یعنی  $f_h$  ها کوچک قابل اغماض باشند، آنگاه

$$Var(\bar{Y}_{st}) = \sum_{h=1}^L w_h^2 \frac{S_h^2}{n_h} = \frac{1}{N^2} \sum_{h=1}^L \frac{N_h^2}{n_h} \cdot S_h^2$$



## فرع ۲

اگر تخصیص متناسب مورد نظر باشد، یعنی

$$n_h = \frac{n}{N} N_h$$

$$\frac{n_h}{N_h} = \frac{n}{N}$$

بنابراین

$$Var(\bar{Y}_{st}) = \sum_{h=1}^L \frac{N_h}{N} \cdot \frac{S_h^2}{n} \left( \frac{N-n}{N} \right) = \frac{1-f}{n} \sum W_h S_h^2$$

## فرع ۳

اگر نمونه گیری باتخصیص متناسب بوده و مقدار واریانس در همه طبقات یکی باشد آنگاه اگر واریانس مشترک را با  $S_W^2$  نشان دهیم تمام  $S_W^2$  ها برابر با  $S_h^2$  هستند بنابراین

$$Var(\bar{Y}_{st}) = \frac{N-n}{N} \cdot \frac{S_w^2}{n} = (1-f) \frac{S_w^2}{n}$$

$f$  کسر نمونه گیری در کل جامعه است

## فرع ۴

اگر  $t_N$  مجموع واحدهای جامعه باشد، آنگاه  $N\bar{Y}_{st}$  برآورد کننده نااریب  $t_N$  است

$$E(\hat{t}_N) = t_N$$

واریانس این برآورد کننده

$$Var(\hat{t}_N) = \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h}$$

## ۳-۶ برآورد واریانس

اگر از هر طبقه نمونه تصادفی ساده بگیریم برآورد کننده نااریب  $S_h^2$ ، بنابراینچه در نمونه گیری تصادفی ساده دیدیم، عبارت است از:

$$S_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (Y_{hi} - \bar{Y}_h)^2$$

## ۳-۷ قضیه

در نمونه گیری تصادفی ساده با طبقه بندی،  
بر آورد کننده نااریب  $Var(\bar{Y}_{st})$  برابر است با



$$\sigma^2(\bar{Y}_{st}) = \hat{Var}(\bar{Y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{s_h^2}{n_h}$$

۳-۸ حدود اطمینان میانگین و مجموع واحدهای جامعه

برای یک نمونه ی تصادفی ساده با طبقه بندی به حجم  
از جامعه ای به حجم  $N$  تعداد نمونه های ممکن

$$\pi_{h=1}^L \begin{bmatrix} N_h \\ n_h \end{bmatrix} n$$

است.



متغیر تصادفی  $\frac{\bar{Y}_{st} - \bar{Y}_N}{\sigma(\bar{Y}_{st})}$  دارای توزیع نرمال استاندارد است. اگر

$z$  مقدار متغیر نرمال استاندارد متناظر با احتمال  $\alpha/2$  باشد به  
سادگی می توان دید که فاصله اطمینان  $100(1 - \alpha)\%$   
برای  $\bar{Y}_N$  بصورت زیر است

فاصله اطمینان تصادفی  $[\bar{Y}_{st} - z\sigma(\bar{Y}_{st}), \bar{Y}_{st} + z\sigma(\bar{Y}_{st})]$

برآورد فاصله اطمینان  $[\bar{Y}_{st} - z\hat{\sigma}(\bar{Y}_{st}), \bar{Y}_{st} + z\hat{\sigma}(\bar{Y}_{st})]$





برآورد فاصله اطمینان برای  $t_N$  با ضریب اطمینان  $1 - \alpha$  بصورت زیر است

$$\left[ N\bar{Y}_{st} - zN\hat{\sigma}(\bar{Y}_{st}), N\bar{Y}_{st} + zN\hat{\sigma}(\bar{Y}_{st}) \right]$$

$$\hat{\sigma}^2(\bar{Y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L g_h s_h^2$$

درجه آزادی موثر با  $ne$  نشان می دهیم که بصورت زیر محاسبه می شود

$$ne = \frac{\left\langle \sum g_h s_h^2 \right\rangle^2}{\sum \frac{g_h^2 s_h^2}{n_h - 1}} ; g_h = \frac{N_h(N_h - n_h)}{n_h}$$

مقدار  $ne$  همیشه بین کوچکترین مقدار از مقادیر  $(n_h - 1)$  و مجموع آنها می افتد. اگر توزیع واحدهای جامعه چاولگی مثبت باشد این فرمول درجه آزادی را بیش برآورد می کند

## ۳-۹ تخصیص اپتیمم

در نمونه گیری با طبقه بندی، مقادیر حجمهای نمونه ای در طبقات مربوط از طرف نمونه گیر تعیین میشوند.....

$$\text{هزینه} = C = C_0 + \sum_{h=1}^L C_h n_h$$

← مبلغ ثابت

ساده ترین تابع هزینه

← هزینه نمونه گیری هر واحد از طبقه  $h$  ام

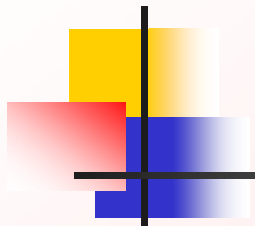
اگر  $C_0$  فقط هزینه رفت و آمد بین طبقات باشد

بهرتر است که بصورت  $\sum t_h \sqrt{n_h}$  نمایش

داده شود که در آن  $t_h$  هزینه رفت و آمدها به

ازای هر واحد است





در نمونه گیری تصادفی با طبقه بندی که تابع هزینه

$$\text{بصورت } C = C_0 + \sum_{h=1}^L C_h n_h \text{ است، برآورد}$$

واریانس  $\bar{Y}_{st}$  برای  $C$  مشخص و هزینه برای  
واریانس مشخص  $Var(\bar{Y}_{st})$  وقتی مینیمم می شود  
که  $n_h$  متناسب با  $W_h S_h / \sqrt{C_h}$  است.



معادله زیر  $n_h$  را بر حسب مقدار  $N_h$  بدست می دهد ولی، مقدار  $n$  رانمی دانیم. مقدار  $n$  در رابطه با این است که  $C$  را از قبل داشته باشیم یا  $V$  را

$$n_h = \frac{W_h S_h / \sqrt{C_h}}{\sum_{h=1}^L (W_h S_h / \sqrt{C_h})} \cdot n = \frac{N_h S_h / \sqrt{C_h}}{\sum_{h=1}^L (N_h S_h / \sqrt{C_h})} \cdot n$$

نتیجه

در طبقه مفروض هر چه ۱- طبقه بزرگتر باشد ۲- ناهمگنی طبقه بیشتر باشد و ۳- مایل به نمونه گیری ارزانتری باشیم باید نمونه ای بزرگتر انتخاب کنیم. زیرا  $n_h$  به  $N_h$  تناسب مستقیم دارد و با  $S_h$  که معرف تغییرات طبقه است نیز تناسب مستقیم دارد ولی با  $C_h$  تناسب معکوس دارد.



الف). وقتی  $C$  معلوم است با معلوم بودن  $N_h$  ها و  $S_h$  ها و  $C_h$  ها ابتدا مقدار  $n$  را می یابیم و سپس مقادیر  $n_h$  بدست می آوریم.

$$n = \frac{(c - c_0) \sum (N_h S_h / \sqrt{c_h})}{\sum (N_h S_h \sqrt{c_h})}$$

ب) وقتی  $V$  از قبل تثبیت شود

وقتی  $V$  معلوم است  $n$  به صورت زیر بدست می آید

$$n = \frac{\left[ \sum_{i=1}^L W_h S_h / \sqrt{c_h} \right] \left[ \sum_{i=1}^L W_h S_h \cdot \sqrt{c_h} \right]}{V + \frac{1}{N} \sum_{h=1}^L W_h S_h^2}$$



**تبصره ۱:** وقتی  $C_h = C$ ، تابع هزینه به صورت

$$C = c_0 + cn$$

در می آید و 
$$n = \frac{C - c_0}{c}$$

$$n_h = \frac{W_h S_h / \sqrt{c}}{\sum_{h=1}^L (W_h S_h / \sqrt{c_h})} \cdot n = \frac{W_h S_h}{\sum_{h=1}^L W_h S_h} \cdot n = \frac{N_h S_h}{\sum_{h=1}^L N_h S_h} \cdot n$$

تخصیص نین

**تبصره ۲:** طبق رابطه بالا  $n_h$  مینیم می شود و مقدار مینیم برابر

$$V = Var(\bar{Y}_{st}) = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h} - \sum_{h=1}^L \frac{W_h S_h^2}{N}$$

$$V_{opt} = \frac{1}{n} \left[ \sum_{h=1}^L W_h S_h \right]^2 - \frac{1}{N} \sum_{h=1}^L W_h S_h^2$$

### ۱۱-۳ مقایسه دقت نمونه گیری تصادفی با طبقه بندی و نمونه گیری تصادفی ساده



### ۱۲-۳ قضیه

اگر  $\frac{1}{N_h}$  قابل اغماض باشد آنگاه  $V_{opt} \leq V_{prop} \leq V_{ran}$

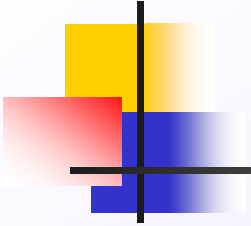
این رابطه نشان می دهد که اگر دو نمونه گیری تصادفی ساده و نمونه گیری تصادفی با طبقه بندی در حالت تخصیص ایتیم را در نظر بگیریم در تقلیل واریانس، دو مولفه دخالت دارند

$$V_{ran} = V_{opt} + \frac{1}{n} \sum_{h=1}^L W_h (s_h - \bar{s})^2 + \frac{1-f}{n} \sum_{h=1}^L W_h (\bar{Y}_h - \bar{Y}_N)^2$$

نتیجه



اگر  $V_{ran} \leq V_{prop}$  نمونه گیری تصادفی ساده کاراتر از نمونه گیری با طبقه بندی و تخصیص متناسب است.



اگر فرض کنیم تمام  $S_h^2$  ها با هم برابرند و مقدار مشترک آنها را با  $S_w^2$  نشان دهیم به شرطی که تخصیص متناسب به مفهوم نیمن، اپتیمم باشد داریم:

$$\sum_{h=1}^L N_h (\bar{Y}_h - \bar{Y}_N)^2 \langle \frac{1}{N} \sum_{n=1}^L (N - N_h) S_w^2$$

$$\text{یا} \sum_{h=1}^L \frac{N_h (\bar{Y}_h - \bar{Y}_N)^2}{L-1} \langle S_w^2$$

یعنی نسبت  $F$  باید کوچکتر از ۱ باشد.

در چنین حالتی نمونه گیری تصادفی کاراتر از نمونه گیری تصادفی با طبقه بندی در حالت تخصیص متناسب است.





### ۳-۱۴ برآورد حجم نمونه برای داده‌های پیوسته

فرض کنیم که  $\bar{Y}^{st}$  دارای واریانس معلوم  $V$  باشد. اگر  $S_h$  برآورد  $S_h$  فرض شود، قرار می‌دهیم  $\frac{n_h}{n} = W_h$  می‌دانیم که :

$$V \approx \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h} - \frac{1}{N} \sum W_h S_h^2$$

مقدار  $n_h$  را جایگزین می‌کنیم داریم

$$V = \frac{1}{n} \sum W_h^2 S_h^2 - \frac{1}{N} \sum W_h S_h^2$$

رابطه کلی برای  $n$

$$\Rightarrow n = \sum \frac{W_h^2 S_h^2}{W_h} / V + \frac{1}{N} \sum W_h S_h^2$$

اگر  $N$  بزرگ باشد

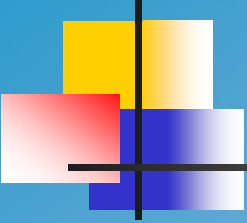
$$\Rightarrow n_o = \frac{1}{V} \sum \frac{W_h^2 S_h^2}{W_h}$$

اگر  $N$  بزرگ نباشد

$$\Rightarrow n = \frac{n_o}{1 + \frac{1}{NV} \sum W_h S_h^2}$$

## حالات خاص


الف) تخصیص را اپتیمم گیریم وقتی


$$n_h = n \frac{W_h S_h}{\sum W_h S_h}$$

آنگاه صورت کسر

یعنی  $\Rightarrow \frac{n_h}{n} = W_h = \frac{W_h S_h}{\sum W_h S_h} \Rightarrow \left( \sum W_h S_h \right)^2$

در نتیجه  $n$  می شود  $\Rightarrow$

$$n = \frac{\left[ \sum_{h=1}^L W_h S_h \right]^2}{V + \frac{1}{N} \sum W_h S_h^2}$$


## حالات خاص

ب) در تخصیص متناسب

$$w_h = \frac{n_h}{n} = \frac{N_h}{N} = W_h$$

$$\Rightarrow n_o = \frac{1}{V} \sum W_h S_h^2$$

$$\Rightarrow n = \frac{n_o}{1 + \frac{n_o}{N}}$$



### ۳-۱۵ آورد مجموع واحدهای جامعه

در این بخش  $V$  معرف  $Var(\hat{t}_N)$  یا  $Var(N\bar{Y}_{st})$  است که برابر  $N^2 Var(\bar{Y}_{st})$  است. از قرار دادن  $\frac{V}{N^2}$  بجای  $V$  در رابطه  $n$  داریم:

$$n = \frac{\sum N_h^2 s_h^2 / w_h}{V + \sum N_h s_h^2}$$

برای تخصیص اپتیمم ( $n$  تثبیت شده) داریم

$$n = \left[ \sum_{h=1}^L N_h s_h \right]^2 / V + \sum_{i=1}^L N_h s_h^2$$

و برای تخصیص متناسب داریم

$$n_o = \frac{N}{V} \sum_{h=1}^L N_h s_h^2 \rightarrow n = \frac{n_o}{1 + \frac{n_o}{N}}$$

## ۳-۱۶ نمونه گیری با طبقه بندی برای نسبت ها


می خواهیم در جامعه ،نسبت واحدهایی را بیابیم که در رده معین  $C$  می افتند

نسبت واحدهایی از طبقه  $h$  ام است که در رده  $C$  می افتد

$$P_h = \frac{A_h}{N_h}$$

نسبت واحدهایی از نمونه طبقه  $h$  ام است که در رده  $C$  می افتد

$$p_h = \frac{a_h}{n_h}$$


$$p_{st} = \sum_{h=1}^L W_h p_h \text{ و } E(p_{st}) = \frac{1}{N} \sum_{h=1}^L A_h = P$$

یعنی  $P_{st}$  برآورد کننده نااریب نسبت واحدهایی از جامعه است که در رده  $C$  هستند.

### ۳-۱۷ قضیه

در نمونه گیری تصادفی با طبقه بندی

$$V(p_{st}) = \frac{1}{N^2} \sum_{h=1}^L \frac{N_h^2 (N_h - n_h)}{N_h - 1} \cdot \frac{P_h Q_h}{n_h}$$

تبصره

در کاربردها، حتی وقتی که  $fpc$  قابل اغماض

نباشد  $\frac{N_h}{N_h - 1}$  را برابر ۱ می گیریم و فرمول

واریانس بصورت تقریبی زیر به کار می رود

$$V(p_{st}) = \sum_{h=1}^L W_h^2 [1 - f_h] \frac{P_h Q_h}{n_h}$$

فرع ۱  
اگر بتوانیم

*fpc* رانادیده بگیریم :

$$V(p_{st}) = \sum W_h^2 \frac{P_h Q_h}{n_h}$$

فرع ۲

اگر تخصیص متناسب باشد

$$\frac{N_h - n_h}{N - n} = \frac{N_h}{N} = \frac{n_h}{n}$$

$$V(p_{st}) = \frac{N - n}{N} \cdot \frac{1}{nN} \sum_{h=1}^L \frac{N_h^2}{N_h - 1} P_h Q_h \approx \frac{1 - f}{n} \sum_{h=1}^L W_h P_h Q_h$$



چون  $P_h$  و  $Q_h$  در عمل مجهولند لذا نمی توان  $V(p_{st})$  را بدست آورد. برای تعیین  $\hat{V}(p_{st})$ ، همانطور که در نمونه گیری تصادفی ساده برای نسبت ها دیدیم بر آورد  $\frac{Q_h P_h}{n_h - 1}$  تقریباً برابر با  $\frac{p_h q_h}{n_h - 1}$  است، لذا

$$\hat{V}(p_{st}) \approx \frac{1}{N^2} \sum_{h=1}^L \frac{N_h^2 (N_h - n_h)}{N_h - 1} \frac{p_h q_h}{n_h - 1}$$





## فرع ۴

بهترین انتخاب  $n_h$  ها برای مینیم کردن  $V(p_{st})$  از قضیه کلی

نتیجه می شود:  
الف) اگر حجم نمونه ثابت باشد برای اینکه واریانس مینیم شود باید

$$n_h \propto N_h S_h \text{ باشد}$$

$$n_h \approx n_o \frac{N_h \sqrt{P_h Q_h}}{\sum_{h=1}^L N_h \sqrt{P_h Q_h}}$$

ب) اگر هزینه  $C = C_o + \sum_{h=1}^L C_h n_h$  ثابت باشد، وقتی واریانس  $p_{st}$

$$n = \frac{(C - c_o) \sum N_h \sqrt{P_h Q_h / C_h}}{\sum N_h \sqrt{P_h Q_h C_h}} \quad \text{مینیم است که داشته باشیم:}$$

### ۳-۱۸ اثر انحراف از تخصیص ایتیم

در این بخش افت دقت نتیجه نمونه گیری را به دلیل عدم توفیق در دستیابی به تخصیص ایتیم مورد بحث قرار می دهیم

$$n'_h = \frac{n(W_h S_h)}{\sum W_h S_h}$$

حجم نمونه در طبقه  $h$  ام

می دانیم  $\Rightarrow$

$$V_{\min}(\bar{Y}_{st}) = \frac{(\sum W_h S_h)^2}{n} - \frac{\sum W_h S_h^2}{N}$$

اگر  $\hat{n}_h$  حجم نمونه ای مورد استفاده در طبقه  $h$  ام باشد مقدار واریانس:

$$V(\bar{Y}_{st}) = \sum \frac{W_h^2 S_h^2}{\hat{n}_h} - \frac{\sum W_h S_h^2}{N}$$

افزایش واریانس ناشی از تخصیص ناکامل عبارت است از

$$V(\bar{Y}_{st}) - V_{\min}(\bar{Y}_{st}) = \sum \frac{W_h^2 S_h^2}{\hat{n}_h} - \frac{1}{n} (\sum W_h S_h)^2$$

داریم

$$W_h S_h = \frac{n_h}{n} \sum W_h S_h$$

باجایگزینی در روابط صفحه قبل:

$$V(\bar{Y}_{st}) - V_{\min}(\bar{Y}_{st}) = \frac{(\sum W_h S_h)^2}{n^2} \sum \frac{[\hat{n}'_h - n'_h]^2}{n'_h}$$

$$\frac{1}{N} \rightarrow 0$$

داریم:

در رابطه واریانس مینیم زمانی که

$$\frac{V_{\min}(\bar{Y}_{st})}{n} = \frac{(\sum W_h S_h)^2}{n^2}$$

از تقسیم دو رابطه اخیر داریم:

$$\frac{V(\bar{Y}_{st}) - V_{\min}(\bar{Y}_{st})}{V_{\min}(\bar{Y}_{st})} = \frac{1}{n} \sum_{h=1}^L \frac{[\hat{n}_h - n'_h]^2}{\hat{n}_h}$$

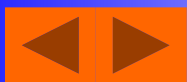
این مقدار، مقدار افزایش نسبی واریانس را وقتی تخصیص بصورت ناکامل انجام می شود نشان می دهد

۳-۱۹ مساله تخصیص با بیش از یک صفت

اِتیِم → 
$$V_{opt} = \frac{(\sum W_h s_h)^2}{n}$$

مِصالِحه ای → 
$$V_{com} = \sum \frac{(W_h s_h)^2}{n_h}$$

مِتناسِب → 
$$V_{prop} = \frac{\sum W_h s_h^2}{n}$$



## ۳-۲۰ ساختن طبقات

### ۳-۲۱ طبقه بندی بعد از انتخاب نمونه

غالباً وقتی نمونه تصادفی ساده دقیقاً بر حسب گروه بندی های جامعه ای حالت تعادلی ندارد، انتخابی مناسب است

میانگین موزون  $\bar{Y}_{st} = \frac{N_1}{N} \bar{y}_1 + \frac{N_2}{N} \bar{y}_2 \quad E(n_h) = nW_h$

$$\hat{V}(\bar{Y}_{st}) = \sum_{h=1}^L W_h^2 \frac{s_h^2}{n_h} - \frac{1}{N} \sum_{h=1}^L W_h s_h^2$$

$$E\left(\frac{1}{n_h}\right) \approx \frac{1}{nW_h} + \frac{1-W_h}{n^2W_h^2}$$

تنها وقتی خوب کار می کند که  $n$  بزرگ و  $n_i$  مثبت باشد

$$\hat{V}_p(\bar{Y}_{st}) = \frac{N-n}{nN} \sum_{h=1}^L W_h s_h^2 + \frac{1}{n^2} \sum_{h=1}^L (1-W_h) s_h^2$$





## فصل چهارم

# نمونه گیری با احتمال

## متغیر

## اهداف رفتاری فصل چهارم

- نمونه گیری تصادفی با احتمال متغیر را تعریف کنید
- نمونه گیری تصادفی متناسب با اندازه را تعریف کنید
- با استفاده از دنباله صفت کمکی و جدول اعداد تصادفی نمونه ای متناسب با احتمال از جامعه انتخاب کنید
- با روش لاهیری مبنی بر استفاده از دنباله صفت کمکی و جدول اعداد تصادفی و مقدار ماکسیمم صفت کمکی، نمونه گیری با احتمال متغیر را انجام دهید



## اهداف رفتاری فصل چهارم

- زوج های موثر و نا موثر و ناموثر را تعریف کنید -
- احتمال موثر یا نا موثر بودن هر زوج را در روش لاهیری بدست آورید
- دلیل روش لاهیری را ارائه دهید
- از روش خرد کردن برای کوچک تر کردن احتمال غیر موثر بودن انتخاب زوج ها در روش لاهیری استفاده کنید.
- ثابت کنید که  
که  
بر آوردن اریب میانگین  $\bar{Z}_n = \sum_{i=1}^n Z_i$  جامعه در نمونه  $Z_i = \frac{Y_i}{NP_i}$  لاهیری با احتمال  
و با جایگذاری است.



## اهداف رفتاری فصل چهارم

- ثابت کنید  $\bar{Z}_n$  برآورد کننده نااریب میانگین جامعه

دارای واریانس  $\frac{\sigma_z^2}{n}$  است  
- ثابت کنید که  $s_z^2$  برآورد کننده نااریب  $\sigma_z^2$  جامعه است

- نمونه گیری با احتمال متغیر وبدون جایگذاری را تعریف کنید



## مقدمه

نوعی از نمونه گیری را که احتمال انتخاب واحدهای جامعه برای شرکت دادن در نمونه از واحدی به واحد دیگر تغییر می کند نمونه گیری با احتمال متغیر می نامند. و در حالت خاصی که احتمالهای انتخاب متناسب با اندازه صفت باشد نمونه گیری تصادفی متناسب با اندازه می گویند و آن را با نماد  $pps$  نشان می دهند.

نمونه گیری تصادفی یا احتمال متغیر به دو روش با  جایگذاری و بدون جایگذاری انجام می شود.

## ۴-۲ شیوه انتخاب نمونه با احتمال متغیر

در نمونه گیری تصادفی ساده از جدول اعداد تصادفی استفاده می کنیم. ولی در این نمونه گیری چون احتمال متناسب به هر واحد با واحدهای دیگر در حالت کلی فرق می کند استفاده مستقیم از جدول اعداد تصادفی میسر نیست. برای انتخاب نمونه اعداد صحیح  $Y_1, Y_2, \dots, Y_N$  را که

متناسب با احتمال های انتخاب واحد ها  $X_1, X_2, \dots, X_N$  جامعه هستند در نظر می گیریم.

صفت کمکی	$Y_1, Y_2, \dots, Y_N$
صفت ملکی	$X_1, X_2, \dots, X_N$

حال به اولین واحد جامعه عدد  $T_1 = X_1$  به دومین واحد جامعه عدد  $T_2 = X_1 + X_2$  و به  $i$  امین واحد جامعه عدد  $T_i = X_1 + X_2 + \dots + X_i$  و به  $N$  امین واحد جامعه عدد  $T_N = \sum_{i=1}^N X_i$  را نسبت می دهیم.



## جدول زیر نتیجه می شود

صفت اصلی	صفت کمکی	$T_i$	$P_i$
$Y_1$	$X_1$	$X_1$	$X_1 / T_N$
$Y_2$	$X_2$	$X_1 + X_2$	
$\vdots$	$\vdots$		
$Y_i$	$X_i$	$X_1 + \dots + X_i$	$X_i / T_N$
$\vdots$	$\vdots$		
$Y_N$	$X_N$	$X_1 + \dots + X_N$	$X_N / T_N$



حال عددی از ۱ تا  $T_N = X_1 + \dots + X_N$  انتخاب و آن را  $R$  می نامیم و با  $T_i$  ها

مقایسه می کنیم و  $T_{i-1} < R \leq T_i$  در این صورت  $Y_i$  ها را به عنوان واحد نمونه

انتخاب کنیم

\* برای اینکه  $T_{i-1} < R \leq T_i$  قرار گیرد تعداد حالات مساعد برابر است با

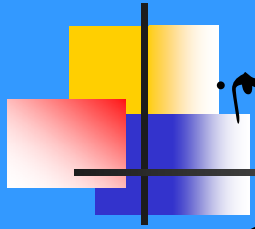
$$T_i - T_{i-1} = (X_1 + X_2 + \dots + X_i) - (X_1 + X_2 + \dots + X_{i-1}) = X_i$$

تعداد حالات ممکن  $R$  برابر با  $T_N$  است پس

$$P(R \text{ وقوع}) = P(Y_i \text{ انتخاب}) = \frac{X_i}{T_N} = \frac{X_i}{X_1 + X_2 + \dots + X_N} = P_i$$



## ۴-۳ روش لاهیری



بزرگترین مقدار صفت کمکی  $X_i$  ها را مشخص و آنرا  $M$  می گیریم.

حجم جامعه  $N$  یک زوج تصادفی  $(i, j)$  انتخاب می کنیم. به قسمی که

$1 \leq i \leq N$  و  $1 \leq M \leq j$  وقتی انتخاب شد  $X_i$  را در نظر می گیریم.

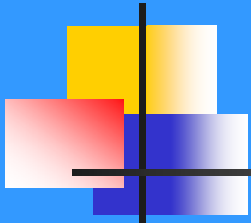
اگر  $X_i \leq j$  آنگاه  $Y_i$  را به عنوان واحدی از نمونه مطلوب اختیار می کنیم.

اگر  $j \geq X_i$  بود زوج  $(i, j)$  را نادیده می گیریم و مجدداً زوج دیگری بر

می گزینیم. این فرآیند آنقدر ادامه پیدا می کند  $n$  تا از  $Y_i$  را انتخاب نماییم.



## ۴-۴ برهان روش لاهیری



زوج های نا موثر=برخی از زوج های  $(i, j)$  که به انتخاب واحد نمونه منجر نمی شود

زوج های موثر= زوجی که به انتخاب واحدی از نمونه منجر می شود

{ابتدا  $i$  از اعداد ۱ تا  $N$  انتخاب شود و سپس داشته باشیم  $j \leq X_i$   $P_1(Y_i) \neq P$ }



چون  $i$  و  $j$  مستقل از هم صورت می گیرند، پس

$$P_1(Y_i) = P(j \leq X_i | i \text{ تا } 1 \text{ تا } N \text{ انتخاب}).$$



$$P(\text{انتخاب } i \text{ از اعداد } 1 \text{ تا } N) = \frac{1}{N}$$

تعداد حالات ممکن انتخاب برابر است زیرا را را از اعداد تا انتخاب کرده ایم.

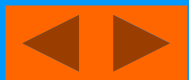
لذا

$$P(j \leq X_i) = \frac{X_i}{M} \quad \Rightarrow \quad P_1(Y_i) = \frac{1}{N} \cdot \frac{X_i}{M}$$

که به طور خلاصه داریم

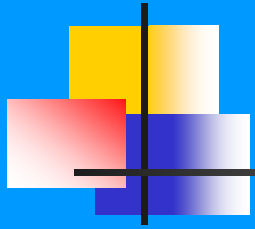
$$P\{(i, j) \text{ غیر موثر بودن زوج}\} = 1 - \frac{\bar{X}_N}{M}$$

$$P(\text{انتخاب } Y_i \text{ برای نمونه}) = \frac{1}{N} \frac{X_i}{M} + \left[ 1 - \frac{\bar{X}_N}{M} \right] \cdot \frac{1}{N} \frac{X_i}{M} + \left[ 1 - \frac{\bar{X}_N}{M} \right]^2 \cdot \frac{1}{N} \frac{X_i}{M} + \dots$$





## در حالت کلی



$$\frac{\bar{X}}{M} < 1 \text{ و قدر مطلق } \frac{\bar{X}_N}{M} - 1 \text{ مقدار کمتر از ۱ است، لذا سری}$$

هندسی بالا همگرا بوده و مجموع جملات آن طبق معمول قابل

محاسبه است. پس وقتی تعداد دفعات انتخابهای متوالی  $(j, j)$  زیاد شود

$$\lim P \left( \text{انتخاب } Y_i \text{ برای نمونه} \right) = \frac{X_i}{T_N}$$

**تبصره:** هر چه  $\frac{\bar{X}_N}{M}$  به یک نزدیکتر باشد احتمال انتخاب غیر موثر کمتر است.

و زمانی که بزرگترین  $X_i$  خیلی با سایر  $X_i$  ها فاصله داشته باشد از روش **خرد کردن** استفاده کنیم

## ۴-۵ روش خرد کردن

روش خرد کردن برای کوچکتر کردن  $1 - \frac{\bar{X}_N}{M}$  ابداع شده

است >> یعنی  $M$  را به چند بخش تقسیم می کنیم <<

ولی در عمل کوچک شدن  $M$  به میزانی انجام می گیرد که

$\frac{\bar{X}_N}{M}$  بزرگ شود و احتمال  $1 - \frac{\bar{X}_N}{M}$  کوچک شود



۴-۶ برآورد میانگین جامعه در نمونه گیری با

احتمال متغیر و با جایگذاری

$$N = \text{حجم جامعه}$$

$$Y_i = \text{واحد } i \text{ ام}$$

$$\sum_{i=1}^N p_i = 1 \quad \text{و} \quad p_i = \text{احتمال انتخاب واحد } i \text{ ام}$$

$$\bar{Z}_n = \frac{1}{n_i} \sum_{i=1}^n Z_i \quad \bar{Z}_i \text{ برآورد کننده نااریب} \quad Z_i = \frac{Y_i}{Np_i}$$

$$\bar{Z}_n = \hat{\bar{Z}}_N = \hat{\bar{Y}}_N$$

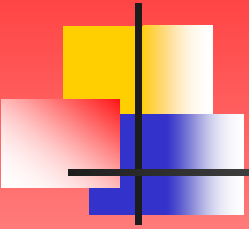


# ۴-۷ واریانس برآورد کننده میانگین جامعه در نمونه گیری با احتمال متغیر و با جایگذاری

$$Var(\hat{\bar{Y}}_N) = Var(\bar{Z}_n) = \frac{1}{n} \left[ \sum_{i=1}^N \frac{Y_i^2}{N^2 p_i} - \bar{Y}_N^2 \right]$$



## تبصره ۱:



اگر  $p_i = \frac{1}{N}$ ، آنگاه نمونه گیری با احتمال  
متغیر به نمونه گیری تصادفی ساده با  
جایگذاری تبدیل میشود. چون  $Z_i = Y_i$

$$Var(\bar{Z}_n) = Var(\bar{Y}_N) = \frac{\sigma_y^2}{n}$$



## تبصره ۲:

اگر احتمال انتخاب  $Y_i$  متناسب با  $Y_i$  باشد آنگاه

$$\frac{Y_i}{\sum_{i=1}^N Y_i} = p_i \quad \text{و} \quad \frac{X_i}{\sum_{i=1}^N X_i} = p_i \Rightarrow Y_i = p_i \sum_{i=1}^N Y_i$$

اگر این مقدار را در رابطه  $Z_i = \frac{Y_i}{Np_i}$  قرار دهیم داریم

$$Z_i = \frac{p_i \sum_{i=1}^N Y_i}{Np_i} = \bar{Y}_N$$



یعنی تمام مقادیر  $Z_i$  با  $\bar{Y}_N$  برابرند و  $\sigma_Z^2 = 0$  این نمونه گیری  
**کارا تر** از نمونه گیری تصادفی ساده با جایگذاری است

### ۳-۸ برآورد واریانس برآورد کننده میانگین در نمونه گیری تصادفی با احتمال متغیر و با جایگذاری

دیدیم که  $\bar{Z}_n$  برآورد کننده نااریب  $\bar{Y}_N$  دارای واریانس  $\frac{\sigma_z^2}{n}$  بود ولی مقدار

واریانس جامعه  $Z_i$  ها مجهول است لذا باید از روی نمونه گیری برآوردی برای این

$$E(s_z^2) = \sigma_z^2$$

واریانس بیابیم.

یعنی  $S_z^2$  برآورد کننده نااریب  $\sigma_z^2$  است. پس با توجه به اینکه

$$Var(\hat{\bar{Y}}_N) = Var(\bar{Z}_n) = \frac{\sigma_z^2}{n}$$

برآورد کننده این واریانس بصورت زیر است که نااریب بودن آن بدیهی است

$$\hat{Var}(\hat{\bar{Y}}_N) = \frac{s_z^2}{n}$$

## نتیجه

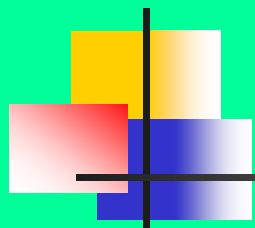
برای استفاده های کاربردی  $V\hat{a}r(\bar{Z}_n)$  می توان  
بصورت زیر نوشت

$$V\hat{a}r(\bar{Z}_n) = \frac{1}{n(n-1)N^2} \sum_{i=1}^n \left[ \frac{Y_i}{p_i} - N\hat{\bar{Y}}_N \right]^2$$



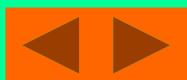


## ۴-۹ نمونه گیری با احتمال متغیر و بدون جایگذاری



### ۴-۱۰ برآوردهای مرتب

ابزاری برای غلبه بر مشکل تغییر امید با هر استخراج، در نظر گرفتن متغیر جدیدی با هر استخراج است به قسمی که امیدش برابر با مقدار جامعه ای متغیر اصلی تحت مطالعه باشد. برآوردهای مبتنی بر چنین ابزارهایی که ترتیب استخراج ها را به حساب می آورند به برآوردهای مرتب موسومند.



## الف) حالت دو استخراج

مقادیر دو واحدی را که در اولین و دومین استخراج به دست آمده اند با  $Y_1$  و  $Y_2$  نشان می دهیم، که این دو واحد الزاماً واحدهای جامعه نیستند و احتمال های اولیه استخراج این دو واحد را به ترتیب  $p_1$  و  $p_2$  می گیریم

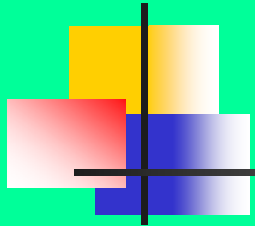
قرار می دهیم

$$Z_1 = \frac{Y_1}{Np_1} \text{ و } Z_2 = \frac{1}{N} \left\{ Y_1 + Y_2 \frac{1-p_1}{p_2} \right\}$$

برآورد کننده نااریب واریانس  $\bar{Z}$  بصورت زیر است

$$n = 2 \text{ وقتی } \hat{Var}(\bar{Z}) = \frac{(1+p_1)^2}{4N^2} \left[ \frac{Y_1}{p_1} - \frac{Y_2}{p_2} \right]^2$$


## ب) حالت کلی



فرض کنید  $Y_n, \dots, Y_2, Y_1$  مقادیر واحدهای باشند که به ترتیب

استخراج شده اند و  $p_n, \dots, p_2, p_1$  احتمال های اولیه انتخاب آنها

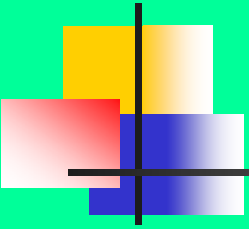
در جامعه باشند متغیر  $Z_i$  را بصورت تعمیمی بیان می کنیم


$$Z_i = \frac{2}{N} \left[ Y_1 + Y_2 + \dots + Y_{i-1} + Y_i \frac{(1 - p_1 - p_2 \dots p_{i-1})}{p_i} \right]$$

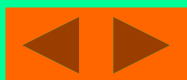
می توان دید که

$$E(Z_i) = \bar{Y}_N \Rightarrow \bar{Z}_n = \frac{1}{n} \sum_{i=1}^n Z_i$$

برآورد ناریبی از  $Var(\bar{Z}_n)$


$$\hat{Var}(\bar{Z}_n) = \bar{Z}_n^2 - \hat{Y}_N^2 = \frac{1}{n(n-1)} \sum_{i=1}^n (Z_i - \bar{Z}_n)^2$$

برآوردهایی که بیان کردیم به ترتیب استخراج واحد ها بستگی داشتند اما متناظر با هر برآورد مرتب، برآوردی نامرتب نیز وجود دارد که به ترتیب استخراج واحد ها بستگی ندارد. معلوم شده که این روش کارا تر از روش برآورد مرتب است، یعنی دارای واریانس کمترند.



پیروز و موفق باشید

پایان

بازگشت به فهرست ها

