Confidence Interval

$$\boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

تخمین بازه‌ای (Interval Estimation)

- $\bar{X}$ و S^2 تخمین‌های نقطه‌ای برای میانگین و واریانس یک جامعه آماری هستند.

- در بسیاری از مواقع ترجیح می‌دهیم به جای این که با استفاده از بردار نمونه‌ها یک نقطه را به عنوان تخمین نقطه‌ای پارامتر θ بدهیم، یک بازه را ارائه کنیم که θ به احتمال زیاد در آن فاصله قرار دارد.

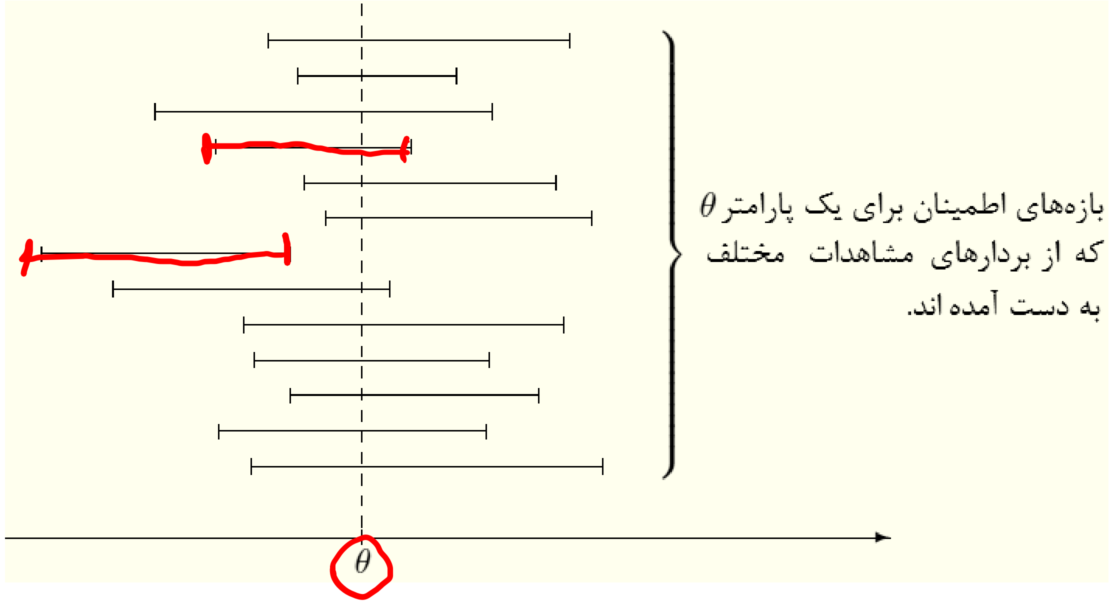
  - مثلاً برای μ، به جای این که یک نقطه $\bar{X}$ را به عنوان تخمین در نظر بگیریم، بازه‌ای را معرفی می‌کنیم که μ به احتمال زیاد داخل آن است.

  - چنین عملی را تخمین بازه‌ای و بازه به دست آمده را یک بازه اطمینان (confidence interval) برای پارامتر θ می‌نامیم.

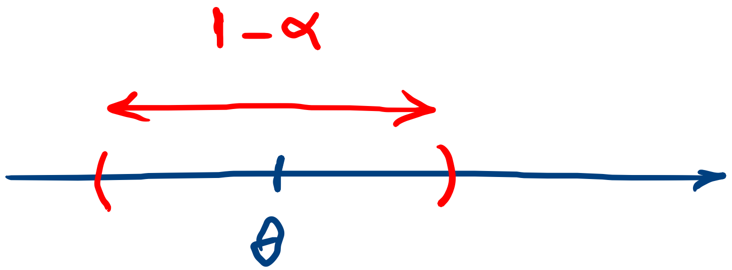
بازه اطمینان (Confidence Interval)

- از آنجا که $\vec{X}$ برداری از متغیرهای تصادفی است، بازه حاصله نیز بازه‌ای تصادفی است که پارامتر θ به احتمال زیادی داخل این فاصله تصادفی قرار دارد.

$\vec{x} = (x_1, \dots, x_n)$



- اگر $P\{a < \theta < b\} = 1 - \alpha$ باشد، بازه (a,b) را بازه اطمینان $1 - \alpha$ و α را سطح اطمینان (Confidence Level) گویند.
  - α معمولا برابر 5%، 1% و یا 0.1% اختیار می‌شود.

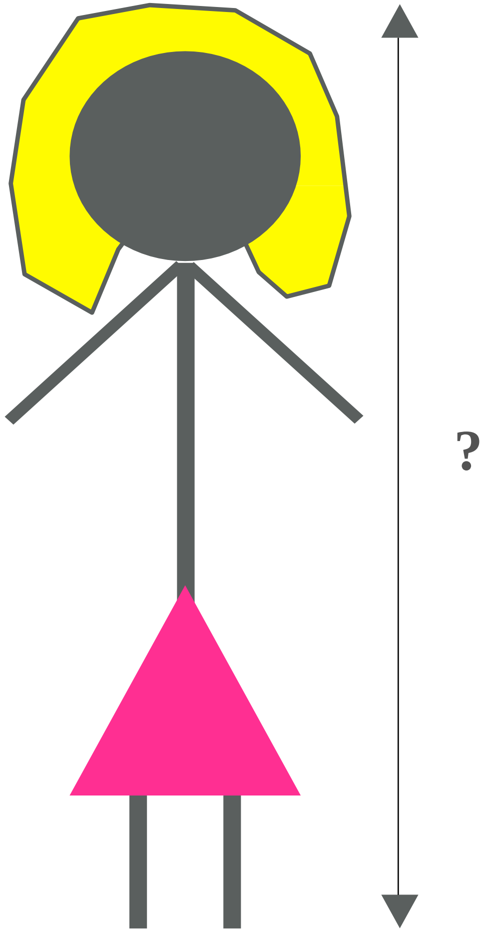


$$P(\underline{a} < \theta < \underline{b}) = \underbrace{1-\alpha}_{\text{confidence level}}$$

بازه اطمینان برای میانگین جامعه

- فرض کنید می‌خواهیم میانگین قد زنان ایرانی را حساب کنیم.

- از آنجا که اندازه‌گیری قد همه زنان ایرانی تقریبا عملی ناممکن است، از نمونه‌برداری کمک می‌گیریم.



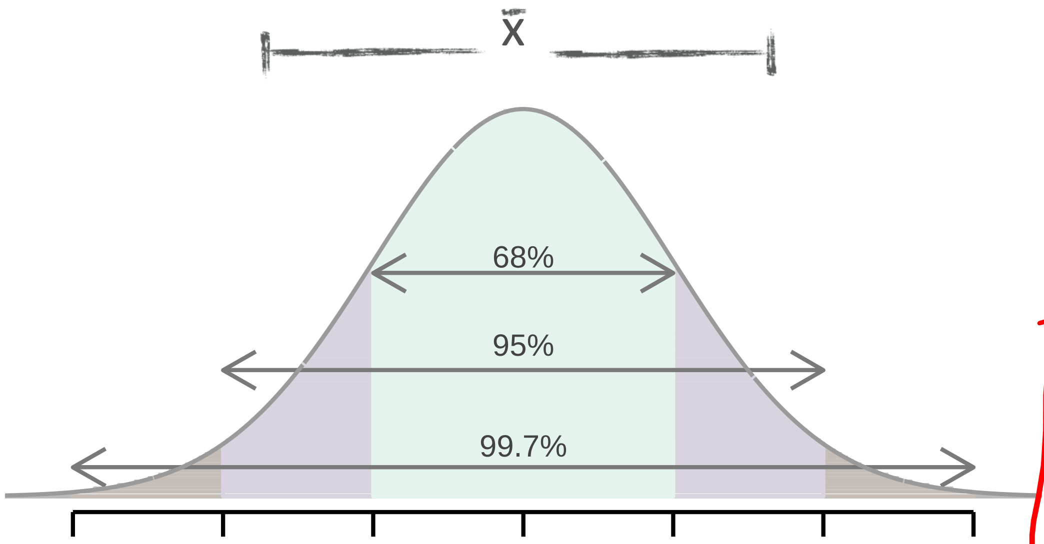
- فرض کنید $\bar{X}$ میانگین نمونه انتخابی ما که اندازه آن n است باشد.
- برای n بزرگ:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

بازه اطمینان ٪۹۵ برای میانگین جامعه

Central Limit Theorem (CLT)

$$\bar{X} \sim N(\mu, \sigma^2/n)$$



$$\boxed{P\left(\mu - 2\frac{\sigma}{\sqrt{n}} < \bar{x} < \mu + 2\frac{\sigma}{\sqrt{n}}\right) \simeq 0.95}$$

$$P\left(-2\frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < 2\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(-2\frac{\sigma}{\sqrt{n}} < \mu - \bar{x} < 2\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$\boxed{P\left(\bar{x} - 2\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 2\frac{\sigma}{\sqrt{n}}\right) = 0.95}$$

بازه اطمینان برای میانگین در صورت معلوم بودن واریانس

- اگر متغیر تصادفی X دارای میانگین نامعلوم μ و واریانس معلوم σ^2 باشد و نمونه‌های X_i به صورت .i.i.d از این متغیر تصادفی برداشته باشیم، تخمین نقطه‌ای زیر را داریم:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \rightarrow \quad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

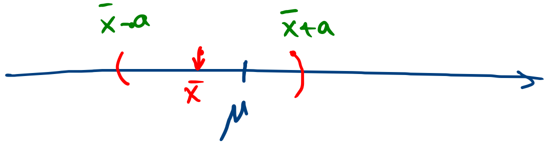
- حال می‌خواهیم یک بازه $(\bar{X} - a,\ \bar{X} + a)$ ارائه دهیم که به احتمال $1 - \alpha$، μ داخل این بازه باشد.

$$\boxed{P(\bar{X}-a<\mu<\bar{X}+a)=1-\alpha}$$

?

$\bar{X}$

$$\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$



$$P(\bar{X}-a<\mu<\bar{X}+a)=P(-a<\mu-\bar{X}<a)=P(-a<\bar{X}-\mu<a)$$

$$=P\left(\underbrace{\frac{-a}{\frac{\sigma}{\sqrt{n}}}}_{-z}<\underbrace{\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}_{Z}<\underbrace{\frac{a}{\sigma/\sqrt{n}}}_{z}\right)$$

$$=P(-z<Z<z)=\Phi(z)-\Phi(-z)=2\Phi(z)-1=1-\alpha$$

$$\Longrightarrow \boxed{\Phi(z)=1-\alpha/2}$$

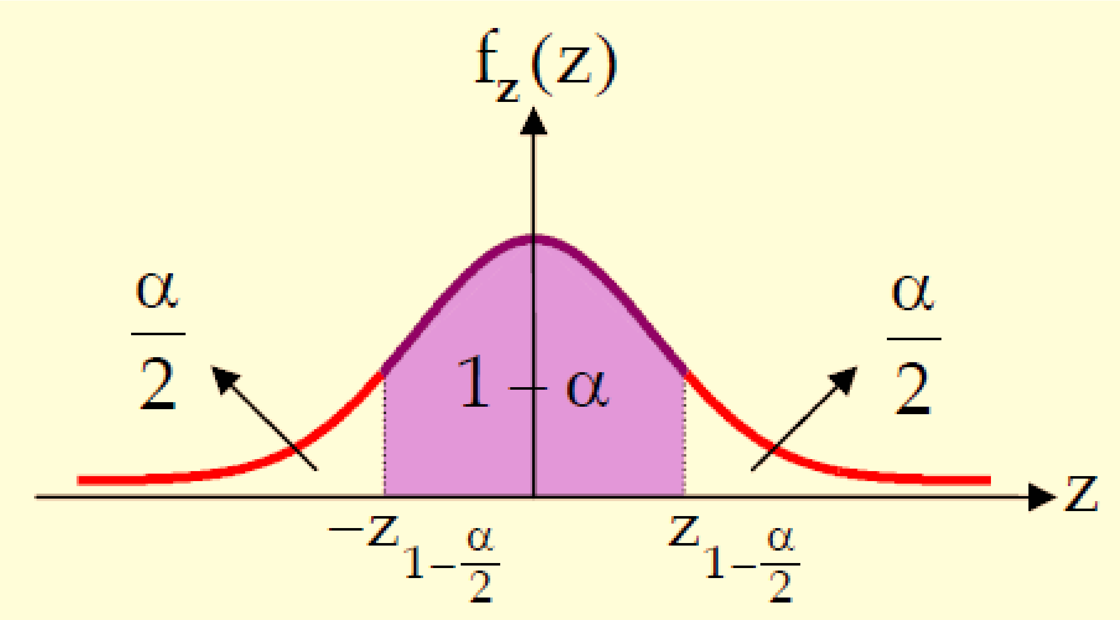
$$\boxed{z=z_{1-\alpha/2}}$$

$$\frac{a}{\frac{\sigma}{\sqrt{n}}}=z_{1-\alpha/2}\Longrightarrow \boxed{a=\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}}$$

$$P(\bar{X}-a<\mu<\bar{X}+a)=P\left(\underline{\bar{X}-\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}}<\mu<\underline{\bar{X}+\frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}}\right)$$

بازه اطمینان برای میانگین با واریانس معلوم

μ



بازه اطمینان برای میانگین با واریانس معلوم

- مثلا برای سطح اطمینان $\alpha = 0.05$ (بازه اطمینان ۹۵٪)، با توجه به جدول داریم:

$$\Phi(1.96) = 0.975$$

$$\mu \in \left(\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{1-\alpha/2} , \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{1-\alpha/2}\right)$$

$$1-\alpha/2 = 0.975 \longrightarrow \boxed{Z_{0.975} = 1.96}$$

بازه اطمینان برای میانگین با واریانس معلوم

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right)$$

مثال ۱

- طول یک محصول دارای توزیع نرمال با انحراف معیار $\sigma = 4mm$ است. اگر در یک نمونه ۳۰ تایی، میانگین نمونه 101mm باشد، بازه اطمینان ۸۰٪ را برای میانگین به دست آورید.

$$\mu \in \left(\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{1-\alpha/2} \, , \, \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{1-\alpha/2} \right)$$

$$1-\alpha = 0.80 \Rightarrow \alpha = 0.2$$

$$Z_{1-\frac{0.2}{2}} = Z_{0.9} = 1.28$$

$$\mu \in \left(101 - \frac{4}{\sqrt{30}} 1.28 \, , \, 101 + \frac{4}{\sqrt{30}} 1.28 \right)$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

ادامه مثال ۱

- پس μ با احتمال ۸۰٪ داخل بازه $(100.07, 101.93)$ قرار دارد.

مثال ۲

- الگوریتم جدیدی را برای محاسبه مدت زمان اجرای آن تست می‌کنیم. فرض کنید می‌دانیم واریانس مدت زمان اجرا $\sigma^2 = 4\,\sec^2$. الگوریتم را n بار اجرا می‌کنیم و هر بار مدت زمان اجرا را اندازه می‌گیریم. فرض کنید متوسط مدت زمان‌های اجرا t ثانیه باشد.

- n حداقل چقدر باشد تا با اطمینان ٪۹۵ بتوان گفت که میانگین حقیقی مدت زمان اجرا (μ) در بازه $t \pm 0.5$ ثانیه است؟

$\bar{x} = t$

$$\mu \in \left(t - \underbrace{\frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}}_{0.5} \, , \, t + \underbrace{\frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}}_{0.5} \right)$$

$$\alpha = 0.05 \longrightarrow z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$$

$$\frac{2}{\sqrt{n}} 1.96 = 0.5 \implies n = \left(\frac{2 \times 1.96}{0.5} \right)^2 = 61.4 \implies \boxed{n = 62}$$

بازه اطمینان برای میانگین با واریانس نامعلوم

- دیدیم که بازه اطمینان $1-\alpha$ برای μ در حالت واریانس معلوم عبارت است از:

$$\left(\bar{X}-\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}},\bar{X}+\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}}\right)$$

- اما در عمل غالباً واریانس جامعه (σ^2) را در اختیار نداریم.

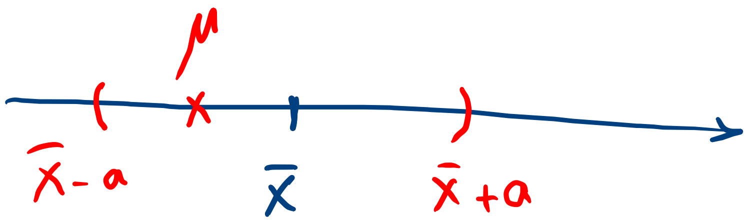
- به این منظور از تخمین نقطه‌ای بی‌غرض σ^2 یعنی واریانس نمونه (S^2) استفاده می‌کنیم:

$$\left(\bar{X}-\frac{S}{\sqrt{n}}z_{1-\frac{\alpha}{2}},\bar{X}+\frac{S}{\sqrt{n}}z_{1-\frac{\alpha}{2}}\right)$$

Confidence Interval:

$$\mu = ? \longrightarrow \bar{X}$$

$$P(\underbrace{\bar{X}-a \le \mu \le \bar{X}+a}_{\text{confidence interval}}) = 1-\alpha \longrightarrow \text{confidence level}$$



$$a = \frac{\sigma}{\sqrt{n}}\, z_{1-\alpha/2}$$

بازه اطمینان برای میانگین با واریانس نامعلوم

- دیدیم که بازه اطمینان $1-\alpha$ برای μ در حالت واریانس معلوم عبارت است از:

$$\boxed{\left(\bar{X}-\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}},\bar{X}+\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}}\right)}$$

- اما در عمل غالباً واریانس جامعه (σ^2) را در اختیار نداریم.

- به این منظور از تخمین نقطه‌ای بی‌غرض $\underline{\sigma}^2$ یعنی واریانس نمونه $(\underline{S}^2)$ استفاده می‌کنیم:

$$\left(\bar{X}-\frac{S}{\sqrt{n}}z_{1-\frac{\alpha}{2}},\bar{X}+\frac{S}{\sqrt{n}}z_{1-\frac{\alpha}{2}}\right)$$

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

شرایط استفاده از CLT برای بازه اطمینان

- جهت استفاده از CLT باید شرایط خاصی برقرار باشند:

۱. شرط استقلال: مشاهداتی که از نمونه‌برداری به دست آمده‌اند باید مستقل از هم باشند.
   - نمونه‌برداری به صورت تصادفی انجام شده باشد.
   - اندازه نمونه از ۱۰٪ کل جامعه آماری کوچکتر باشد.

۲. شرط اندازه نمونه: هر چه اندازه نمونه بزرگتر باشد، استفاده از قضیه CLT معقولتر خواهد بود.
   - اندازه نمونه حداقل ۳۰ باشد.
   - هر چقدر چولگی بیشتر باشد (تقارن کمتری داشته باشد)، اندازه نمونه بزرگتری لازم است.

بازه اطمینان برای نسبت (proportion)

- اغلب مواردی پیش می آید که لازم است نسبت خاصی را در جامعه برآورد کنیم.
  - نسبت افراد بیکار بالای ۱۸ سال در ایران
  - نسبت تعداد دانشجویان دختر دانشگاه تهران
  - نسبت افرادی که در انتخابات به یک فرد خاص رای می‌دهند

- معمولاً نسبت را با p نمایش می‌دهیم:

$$p = \frac{X}{N}$$

که N اندازه کل جامعه، و X تعداد افراد دارای خصوصیت مورد نظر است.

- تخمین p که یک پارامتر جامعه است را با $\hat{p}$ نمایش می‌دهیم.

بازه اطمینان برای نسبت (proportion)

- توجه کنید که می‌توانیم p را به صورت

$$p = \frac{X_1 + X_2 + \ ... + X_N}{N}$$

که X_i متغیر شاخص مربوط به فرد i-ام از جامعه است. به عبارت دیگر $X_i = 1$، اگر فرد i-ام دارای ویژگی مورد نظر باشد، و $X_i = 0$ در غیر این صورت.

- پس می‌توان گفت که p در واقع میانگین یک توزیع برنولی است و قصد ما پیدا کردن یک بازه اطمینان برای p است.

- مشابه با حالت محاسبه بازه اطمینان برای میانگین، می‌توانیم از نمونه‌برداری کمک بگیریم.

نمونه‌برداری جهت تخمین نسبت

- فرض کنید نمونه‌ای به اندازه n از جامعه انتخاب شده است: $X_1, X_2, \dots, X_n$

- دیدیم که بهترین تخمین برای p (که میانگین جامعه است) برابر است با:

$$\hat{p} = \bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- طبق قضیه حد مرکزی داریم:

$$\hat{p} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\boxed{\mu = p} \qquad \boxed{\sigma^2 = p(1-p)}$$

$$\boxed{\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)}$$

s^2

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \le \mu \le \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}$$

$$\hat{p} - \sqrt{\frac{p(1-p)}{n}}\, z_{1-\alpha/2} \le p \le \hat{p} + \sqrt{\frac{p(1-p)}{n}}\, z_{1-\alpha/2}$$

$$\boxed{\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\, z_{1-\alpha/2} \le p \le \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\, z_{1-\alpha/2}}$$

بازه اطمینان برای نسبت

- مشابه با بازه اطمینان $(1-\alpha)$ درصد برای میانگین داریم:

$$\hat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

- توجه کنید که فرض معلوم بودن واریانس جامعه در اینجا برقرار نیست، زیرا معلوم بودن واریانس ($p(1-p)$) معادل با معلوم بودن نسبت p است!

- دو راه برای حل این مشکل وجود دارد:
  - در نظر گرفتن بزرگترین واریانس ممکن که به ازای $p = 0.5$ اتفاق میافتد.
  - استفاده از تخمین نقطه‌ای $\hat{p}$ به جای p.

$$p(1-p) = p - p^2 \longrightarrow 1-2p=0 \Rightarrow p = 1/2$$

بازه اطمینان برای نسبت

- راه دوم:

$$\hat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

شرایط استفاده از CLT برای بازه اطمینان

- جهت استفاده از CLT برای بازه اطمینان نسبت یک جامعه، باید شرایط خاصی برقرار باشند:

۱. شرط استقلال: مشاهداتی که از نمونه‌برداری به دست آمده‌اند باید مستقل از هم باشند.
   - نمونه‌برداری به صورت تصادفی انجام شده باشد.
   - اندازه نمونه از ۱۰٪ کل جامعه آماری کوچکتر باشد.

۲. شرط اندازه نمونه: اندازه نمونه باید به قدری بزرگ باشد که np و $n(1-p)$ هر دو بزرگتر از ۱۰ باشند:

$n > 30$

$$np > 10 \quad \text{and} \quad n(1-p) > 10$$

مثال ۱

- فرض کنید محموله بزرگی از یک کالای خاص در اختیار داریم. از آنجا که بررسی همه این محموله نیازمند زمان و هزینه بالایی است، تنها به بررسی یک نمونه ۲۰۰ تایی می‌پردازیم. ۲۴ مورد از کالاهای بررسی شده خراب تشخیص داده می‌شوند. بازه اطمینان ۹۶٪ برای نسبت کالاهای خراب را پیدا کنید.

$1-\alpha = 0.96 \Rightarrow \alpha = 0.04$ $\qquad 1-\alpha/2 = 0.98 \longrightarrow Z_{0.98} = 2.06$

$\hat{p} = \frac{24}{200} = 0.12$

$0.12 - \sqrt{\frac{0.12(1-0.12)}{200}}\ 2.06 \le p \le 0.12 +$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

ادامه مثال ۱

$$z_{1-\frac{\alpha}{2}} = 2.06$$

$$\hat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.12 - 2.06\sqrt{\frac{0.12 \times 0.88}{200}} < p < 0.12 + 2.06\sqrt{\frac{0.12 \times 0.88}{200}}$$

96% confidence interval for p: (0.1195, 0.1205)

مثال ۲

- یک موسسه آمار و نظرسنجی، می‌خواهد درصد افرادی را که به یک کاندیدای خاص در انتخابات آینده ریاست جمهوری رای می‌دهند (p) تخمین بزند. کمترین تعداد افرادی که نیاز است از آنها نظرسنجی شود تا با اطمینان ۹۵٪ بتوانیم بگوییم که تخمین $\hat{p}$ حداکثر $\pm 3\%$ خطا دارد، چقدر است؟

$$\hat{p} - \underset{0.03}{a} \le p \le \hat{p} + a$$

$$\boxed{n = ?}$$

$$\hat{p} - z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le p \le \hat{p} + \underbrace{z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}}_{0.03}$$

$p-0.03$ $\quad p \quad$ $p+0.03$

$$z_{1-\alpha/2}\sqrt{\frac{\frac{1}{2}(1-1/2)}{n}} \leq 0.03$$

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow 1-\alpha/2 = 0.975 \Rightarrow z_{0.975} = 1.96$$

$$\frac{1.96}{2}\frac{1}{\sqrt{n}} \leq 0.03 \Rightarrow n \geq \left(\frac{0.98}{0.03}\right)^2 \Rightarrow \boxed{n = 1068}$$