Interval Estimation

Bruce M. Boghosian

Department of Mathematics

Tufts University

Outline

Motivation

Bruce M. Boghosian

- Point estimates (MLE or MM) yield a single result.
- There is no indication of how accurate that result is.
- We need a way to quantify the level of uncertainty in the result.
- This is done by constructing a *confidence interval*.
- A confidence interval is an interval in which the parameter has a high probability of being found.
- For example, a *95% confidence interval* for parameter p is an interval surrounding the estimate constructed so that the probability that the actual value of p is in the interval is 95%.

Review of moment-generating functions I

- *Moment-generating function* for continuous r.v. X,

$$M_X(t) := E\left(e^{tx}\right) = \int dx\ f_X(x) e^{tx} = \int dx\ f_X(x) \sum_{j=0}^{\infty} \frac{t^j x^j}{j!} = \sum_{j=0}^{\infty} \frac{t^j}{j!} E\left(x^j\right)$$

$$= 1 + \frac{t}{1!} E(x) + \frac{t^2}{2!} E(x^2) + \frac{t^3}{3!} E(x^3) + \frac{t^4}{4!} E(x^4) + \cdots$$

- The above makes it clear that

$$M_X(0) = 1, \quad M'_X(0) = E(X), \quad M''_X(0) = E(X^2), \quad \cdots \quad \boxed{M^{(k)}_X(0) = E(X^k).}$$

Review of moment-generating functions II

- Suppose that $Y = aX + b$ is a new r.v. linearly related to X,

$$M_Y(t) := E\left(e^{ty}\right) = E\left(e^{t(ax+b)}\right) = E\left(e^{atx}e^{tb}\right) = e^{tb}M_X(at).$$

- Suppose that X_1 and X_2 are uncorrelated, and $Y = X_1 + X_2$,

$$\begin{aligned} M_Y(t) =& E\left(e^{ty}\right) = E\left(e^{t(x_1+x_2)}\right) \\ =& E\left(e^{tx_1}e^{tx_2}\right) = E\left(e^{tx_1}\right)E\left(e^{tx_2}\right) = M_{X_1}(t)M_{X_2}(t) \end{aligned}$$

- The generalization if $Y = X_1 + X_2 + \cdots + X_n$ is then

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t).$$

Review of moment-generating functions III

- Moment generating function of a standard normal Z

$$
\begin{aligned}
M_Z(t) =& E\left(e^{tz}\right)\\
=& \int dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) e^{tz}\\
=& \int dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2-2zt+t^2}{2}+\frac{t^2}{2}\right)\\
=& \exp\left(\frac{t^2}{2}\right)\frac{1}{\sqrt{2\pi}}\int dz\ \exp\left[-\frac{(z-t)^2}{2}\right]\\
=& e^{t^2/2}.
\end{aligned}
$$

Outline of proof of Central Limit Theorem I

- Suppose $X_1, X_2, \ldots, X_n$ are *independent, identically distributed* random variables with mean μ and variance σ.
- Define the random variable

$$Z := \frac{\frac{1}{n}\sum_{j=1}^{n} X_j - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{j=1}^{n} X_j - n\mu}{\sqrt{n}\,\sigma} = \sum_{j=1}^{n} \frac{X_j - \mu}{\sqrt{n}\,\sigma} = \sum_{j=1}^{n} \frac{S_j}{\sqrt{n}}$$

- Where the identically distributed S_j are defined by $S_j := \frac{X_j - \mu}{\sigma}$, so that
  - $E(S_j) = E\left(\frac{X_j-\mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(X_j) - \mu\right] = 0$ (zero mean)
  - $\text{Var}(S_j) = \frac{1}{\sigma^2}\text{Var}\left((X_j - \mu)^2\right) = 1$ (unit variance)

Outline of proof of Central Limit Theorem II

Bruce M. Boghosian

- We have

$$Z = \sum_{j=1}^{n} \frac{S_j}{\sqrt{n}}$$

where $S_j := \frac{X_j - \mu}{\sigma}$ are i.i.d., with zero mean and unit variance

- Let M_S be the moment generating function of (all) the S_j, so

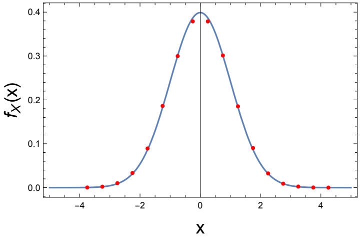
$$\begin{aligned}\lim_{n\to\infty} M_Z = \left[M_S\left(\frac{t}{\sqrt{n}}\right)\right]^n &= \lim_{n\to\infty}\left[1 + \frac{t}{\sqrt{n}}0 + \frac{t^2}{2n}1 + \cdots\right]^n \\ &= \lim_{n\to\infty}\left(1 + \frac{t^2}{2n} + \cdots\right)^n = e^{t^2/2}\end{aligned}$$

- $\boxed{\text{So } Z := \frac{\frac{1}{n}\sum_{j=1}^{n} X_j - \mu}{\sigma/\sqrt{n}} \text{ must be distributed like a standard normal.}}$

Example of the Central Limit Theorem in action

Bruce M. Boghosian

- Random number generator produces uniformly distributed $X_j \in [0, 1]$.
- Each of these has mean $1/2$ and standard deviation $\sigma = 1/2\sqrt{3}$.
- Define $S_j := \frac{X_j - 1/2}{1/(2\sqrt{3})} = \sqrt{3}\,(2X_j - 1)$
- Choose a large number n, and define $Z := \frac{1}{\sqrt{n}}\,(S_1 + S_2 + \cdots + S_n)$
- Do this many times and histogram the results.
- For $n = 20$ and 10,000 histogrammed results:



Example illustrating general methodology

- Suppose that you have n measurements of experimental data.
- You have a priori knowledge that each datum is distributed normally.
- You know the variance $v_0 = \sigma_0^2$, but you do not know the mean μ.

$$p_X(x) = \frac{1}{\sqrt{2\pi}\ \sigma_0} \exp\left[-\frac{(x-\mu)^2}{2\sigma_0^2}\right]$$

- The MLE for the mean is the sample mean, $\mu_e = \frac{1}{n}\sum_{j=1}^{n} x_j$.
- Hence $Z = \frac{\mu_e - \mu}{\sigma_0/\sqrt{n}}$ must be distributed like a standard normal

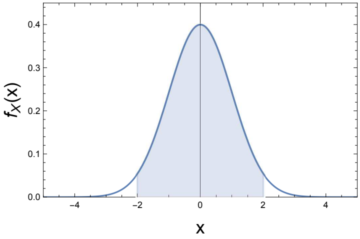
Calculation of confidence intervals

Bruce M. Boghosian

- We know $Z = \frac{\mu_e - \mu}{\sigma_0/\sqrt{n}}$ is distributed like a standard normal.
- Solution to

$$\int_{-z}^{+z} dx \, \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) = 0.95$$

is $z = 1.9599\ldots$.



Calculation of confidence intervals

- With 95% confidence, $Z \in [-z, +z]$ where $z = 1.9599\ldots$
- Hence with 95% confidence, we have

$$-z \le \frac{\mu_e - \mu}{\sigma_0/\sqrt{n}} \le +z$$
$$-z\frac{\sigma_0}{\sqrt{n}} \le \mu_e - \mu \le +z\frac{\sigma_0}{\sqrt{n}}$$
$$\mu_e - z\frac{\sigma_0}{\sqrt{n}} \le \mu \le \mu_e + z\frac{\sigma_0}{\sqrt{n}}$$

- Confidence interval is

$$\boxed{\mu \in \left[\mu_e - 1.9599\frac{\sigma_0}{\sqrt{n}}, \mu_e + 1.9599\frac{\sigma_0}{\sqrt{n}}\right] \quad \text{with a confidence of 95\%}}$$

Z tables

- To make this calculation for any confidence level c, we need to be able to solve

$$\int_{-z}^{+z} dx \, \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) = \alpha$$

- For $\alpha = 0.95$, we know $z = 1.9599\ldots$, but what about other values of c?
- Table A.1 of your text, contains a detailed table of the values of

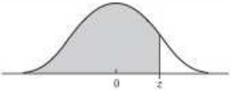
$$\Phi(z) = \int_{-\infty}^{z} dx \, \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

- To find the 95% confidence interval, note that the area of the unshaded region would be $\alpha/2 = 0.025$ on each side.
- From the table, we see $\Phi(+1.96) = 1 - 0.025 = 0.975$, and $\Phi(-1.96) = 0.025$.
- Note that 1.96 is very close to the (more accurate) value of 1.9599 used above.

Z tables (from Larsen & Marx)

Bruce M. Boghosian

Table A.1 Cumulative Areas under the Standard Normal Distribution



z	0	1	2	3	4	5	6	7	8	9
−3.	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0126	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0238	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0300	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0570	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867

Z tables (from Larsen & Marx)

Table A.1 Cumulative Areas under the Standard Normal Distribution (cont.)

z	0	1	2	3	4	5	6	7	8	9
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	0.7703	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	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	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	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	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	0.9874	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
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981

Z tables

- More generally, we write $z_{\alpha/2}$ to be the value such that

$$\text{Prob}(Z \geq z_{\alpha/2}) = \frac{\alpha}{2}.$$

- In our case, if $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$.
- A $100(1-\alpha)\%$ confidence interval for μ_e is then

$$\left[\mu_e - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}, \mu_e + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right]$$

Confidence intervals for the binomial parameter p

Bruce M. Boghosian

- Suppose we conduct n Bernoulli trials with heads probability p.
- For one trial, the mean is p and the standard deviation is $\sqrt{p(1-p)}$
- For n trials, we have a binomial probability distribution with mean p and standard deviation $\sqrt{p(1-p)/n}$
- Using MLE or MM, we have $p_e = \frac{1}{n}\sum_{j=1}^{n} k_j$, so for large n

$$Z = \frac{p_e - p}{\sqrt{p_e(1-p_e)/n}}$$

will be distributed like a standard normal, by the Central Limit Theorem.

Margin of error

- The *margin of error* is half the maximum width of a 95% confidence interval.
- Let k be the number of successes in n Bernoulli trials. Estimate is $p_e = k/n$.
- The margin of error is $100d\%$ where $d = \frac{1.9599}{2\sqrt{n}}$
- The definition can be generalized to values of α other than 0.05.
- The generalization is $100d\%$ where $d = \frac{z_{\alpha/2}}{2\sqrt{n}}$.

Choosing sample sizes

Bruce M. Boghosian

- We have in general

$$\text{Prob}\left(-d \leq \frac{X}{n} - p \leq +d\right) = 1 - \alpha.$$

- This can be regarded as an equation for the minimum value of n needed to attain the confidence α.

Summary

Bruce M. Boghosian

- We have reviewed moment-generating functions
- We have reviewed the spirit and proof of the Central Limit Theorem
- We have defined and shown how to calculate confidence intervals
- We have learned how to read the z tables in Appendix A.1
- We have learned how to generalize this to Bernoulli trials
- We have defined margin of error, and estimated needed sample sizes