Ludwig-Maximilians-Universität München
Institut für Informatik
Prof. Dr. Matthias Schubert
Felix Borutta

Munich, May 29, 2019

Managing Massive Multiplayer Online Games

SS 2019

Exercise Sheet 5: Game Analytics

The assignments are due June 05, 2019

Assignment 5-1 *Bot Detection with Bayes*

Consider an abstract game in which players regularly have to make decisions (e.g., whether to go south, north, east or west). We assume that there are always four alternatives $\{a_1, ..., a_4\}$ of such actions and that a BOT selects each of these with the same probability. With the help of log data it could be estimated empirically that human players select the alternatives with the following probabilities:

$$P(a_1) = 10\%, P(a_2) = 20\%, P(a_3) = 30\%, P(a_4) = 40\%.$$

Player p_1 was observed to have the following sequence of decisions:

$$O = [a_3, a_2, a_1, a_4, a_1, a_2, a_2, a_3, a_1]$$

In the following B is the event that player p_1 is a BOT and $\overline{B}$ is the event that player p_1 is a human player.

(a) Calculate the probability $P(O \mid B)$ that a BOT produces the given sequence.

$$P(O|B) = 0.25^9 = 3.8 \cdot 10^{-6}$$

(b) Calculate the probability $P(O \mid \overline{B})$ that a human player produces the given sequence.

$$P(O|\overline{B}) = 0.1^3 \cdot 0.2^3 \cdot 0.3^2 \cdot 0.4 = 2.88 \cdot 10^{-7}$$

Note: for the sake of simplicity, we assumed the actions to be performed independent from each other here. Otherwise we would have to deal with conditional probabilities as follows

$$\begin{aligned} P(O|\overline{B}) &= P(action = a_3) \cdot P(action = a_2|[a_3]) \cdot P(action = a_1|[a_3, a_2]) \cdot P(action = a_4|[a_3, a_2, a_1]) \\ &\cdot \ldots \cdot P(action = a_1|[a_3, a_2, a_1, a_4, a_1, a_2, a_2, a_3]) \end{aligned}$$

(c) Assume that 1% of all players are BOTs. Calculate the probability $P(B \mid O)$ that player p_1 is a BOT.

$$P(B|O) = \frac{P(O|B) \cdot P(B)}{P(O)} =$$

Law of total probability:

$$\frac{P(O|B) \cdot P(B)}{P(O|B) \cdot P(B) + P(O|\overline{B}) \cdot P(\overline{B})} =$$

$$\frac{3.8 \cdot 10^{-6} \cdot 10^{-2}}{3.8 \cdot 10^{-6} \cdot 0.01 + 2.88 \cdot 10^{-7} \cdot 0.99}$$
$$= 0.1176$$

Assignment 5-2 *Probabilistic Balancing*

Consider another game where players can choose between several different settings (e.g. races, classes, fractions) in the beginning. Let $s_1, ..., s_n$ denote such settings.

Assume that 1000 games between players with settings s_1 and players with settings s_2 have been recorded. 400 of those were won by the players having settings s_1.

Are the settings s_1 and s_2 well balanced? Assume a significance level of $\alpha = 0.05$ to confirm or reject the hypothesis. Calculate the probability of this observation assuming that the game is fair, i.e., that the chances for winning is equal for both players.

Calculate the probability of this observation assuming that the game is fair:

$$\begin{aligned}
P(B(N,p) = i) &= \binom{N}{i} \cdot p^i \cdot (1-p)^{N-i} \\
P(B(1000, 0.5) = 400) &= \binom{1000}{400} \cdot 0.5^{400} \cdot (1-0.5)^{1000-400} \\
P(B(1000, 0.5) = 400) &= 4.6339 \cdot 10^{-11}
\end{aligned}$$

Are the settings s_1 and s_2 well balanced?

Solution 1. Left-tailed testing:

null hypothesis: $H_0 : p = 0.5$, alternative hypothesis $H_A : p < 0.5$, i.e., we assume from our observation that the probability for players with setting s_1 to win is less than 50%.

The sample proportion is: $\hat{p} = \frac{400}{1000} = 0.4$

The test statistic, resp. the Z-value is, therefore:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot (1-p)}{N}}} = \frac{0.4 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{1000}}} = -6.32$$

Given that the Z-value for our significance level $\alpha = 0.05$ is $Z_\alpha = -1.645$[1], we can reject the null hypothesis. There is enough evidence for $\alpha = 0.5$ that we can say that the win probability for players with setting s_1 is less than 50%.

Another variant is to approximate the binomial distribution explicitly with the normal distribution, which is possible due to fulfilling the Laplace theorem $\sqrt{N \cdot p \cdot (1-p)} = \sqrt{1000 \cdot 0.5 \cdot 0.5} = 15.81 > 3$, is as follows:

[1] You can check these values in the quantile table for the normal distribution as the standard deviation fulfills the Laplace theorem, i.e., $\sigma > 3$, resp. $\sqrt{N \cdot p \cdot (1-p)} > 3$.

$$\begin{aligned}
P(B(N,p) \leq i) &= \sum_{j=0}^{i} P(B(n,p) = j) \\
B(1000, 0.5) &\approx N(1000 \cdot 0.5, 1000 \cdot 0.5 \cdot 0.5) = N(500, 250) \\
P(N(500,250) \leq 400) &= P(500 + N(0,250) \leq 400) \\
&= P(N(0,250) \leq -100) \\
&= P(N(0,1) \cdot \sqrt{250} \leq -100) \\
&= P(N(0,1) \leq \frac{-100}{\sqrt{250}}) \\
&= P(N(0,1) \leq -6.32) = 8.5 \cdot 10^{-8}
\end{aligned}$$

Note that the last step is a read-out from the quantile table. The result is a probability less than the significance level $\alpha = 0.05$.

Solution 2. two-tailed testing:

null hypothesis: $H_0 : p = 0.5$, alternative hypothesis $H_A : p \neq 0.5$, i.e., the game is not fair.

Assuming H_0, the expected value for the outcome should be

$$\mu = 0.5 \cdot 1000 = 500$$

with a standard deviation of

$$\sigma = \sqrt{N \cdot p \cdot (1-p)} = \sqrt{1000 \cdot 0.5 \cdot 0.5} = 15.81.$$

Since we are doing a two-tailed test here, we define the bounds of our acceptance range with respect to

$$z_{\frac{\alpha}{2}} = 1.96.$$

This value is taken from the quantiles table of the normal distribution (again possible due to fulfilling the Laplace theorem).

Given those values we can calculate our acceptance range as

$$\begin{aligned}
A &= [\mu - z_{\frac{\alpha}{2}} \cdot \sigma; \mu + z_{\frac{\alpha}{2}} \cdot \sigma] \\
&= [470; 530],
\end{aligned}$$

and derive the rejection area

$$\overline{A} = [0; 469] \cup [530; 1000].$$

The value 400 falls into the rejection area and thus the null hypothesis can be rejected.

We also can compute the Error of Type 2[2], i.e., the probability to accept the hypothesis although it is wrong. Note that we use the probability value from our observation, hence $p = 0.4$, and the amount of wins from our null hypothesis, hence $A = [470; 530]$! In other words, what is the probability to end up in the acceptance range when using the win probability from our observation?

$$\begin{aligned}
P(a \leq X \leq b) &\approx \Phi(\frac{b + 0.5 - Np}{\sqrt{N \cdot p \cdot (1-p)}}) - \Phi(\frac{a + 0.5 - Np}{\sqrt{N \cdot p \cdot (1-p)}}) \\
P(470 \leq X \leq 530) &\approx \Phi(\frac{530 + 0.5 - 400}{\sqrt{1000 \cdot 0.4 \cdot 0.6}}) - \Phi(\frac{470 + 0.5 - 400}{\sqrt{1000 \cdot 0.4 \cdot 0.6}}) \\
&= \Phi(\frac{130.5}{\sqrt{240}}) - \Phi(\frac{70.5}{\sqrt{240}}) \\
&= \Phi(8.42) - \Phi(4.55) \\
&\approx 0
\end{aligned}$$

[2]The Error of Type 1 is given by the significance value α

With a probability very close to 0% the null hypothesis would be assumed to be true although it's wrong.

Quantiles Table for the normal distribution:

How to read from quantiles table: The left far column are the first and second digits of the z-value, the values in the very first row are the third digits. The values inside the body of the table are the quantiles. So if we have some significance value of let's say 2.5%, how can we get the corresponding z-value? The first thing we can see is that the table only contains values $0.5 \leq x < 1$, which corresponds to everything that is on the right side of the mean (0.5 is the mean of the standard normal distribution). If we now have a value of 2.5%, we check the corresponding z-value for $1 - 0.025 = 0.975$, which is possible since the distribution function is symmetric. So given 0.975, we search this value in the table and see that it is in the row denoted with "1.9" and column denoted as "0.06". This means that the z-value for 0.975 is 1.96, and since we are actually looking for the z-value of $1 - 0.975$ we need to go "on the other side of the mean" and therefore we take -1.96 as z-value for 0.025.

	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
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
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Abbildung 1: Quantiles Table for the normal distribution.