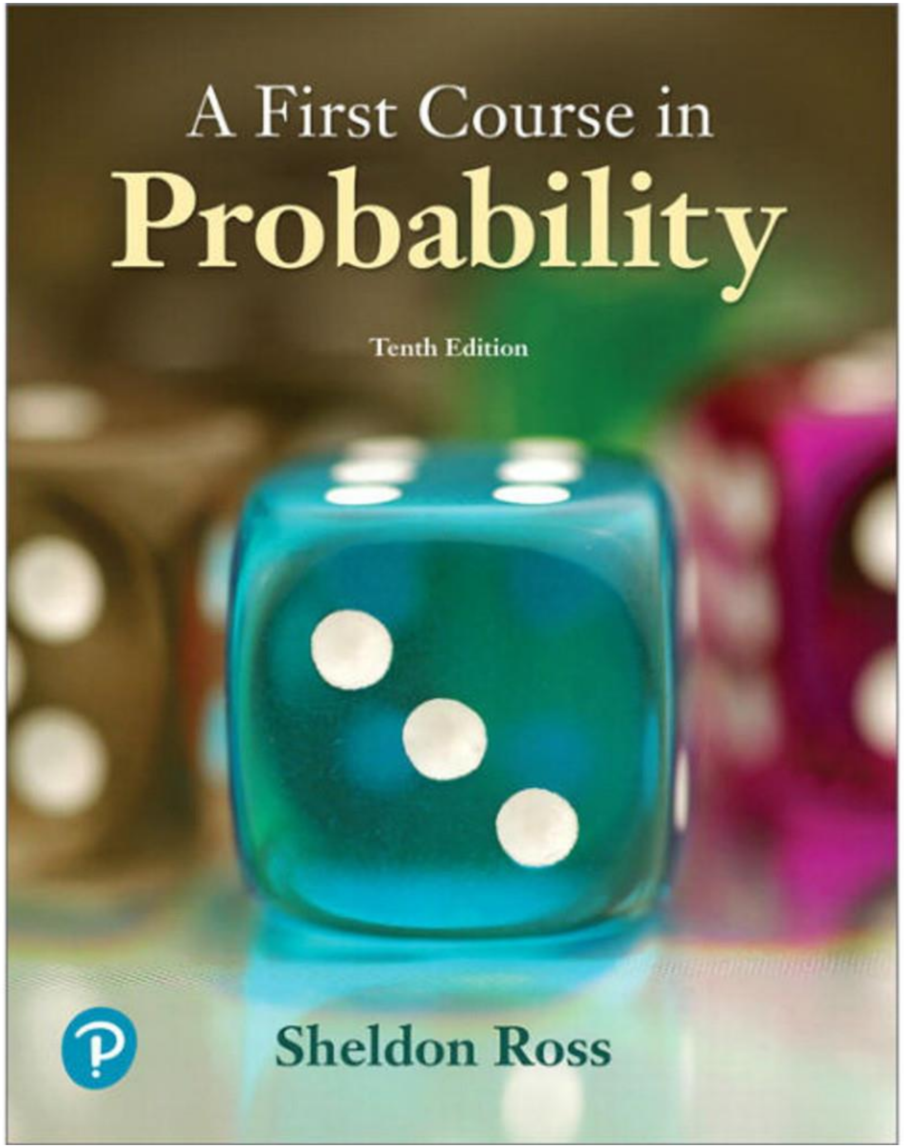
Lecture 9.3

Normal approximation to binomial

PURDUE UNIVERSITY | Department of Mathematics



Today's reading: 5.4.1

Next class: 5.5

HW7 due now!

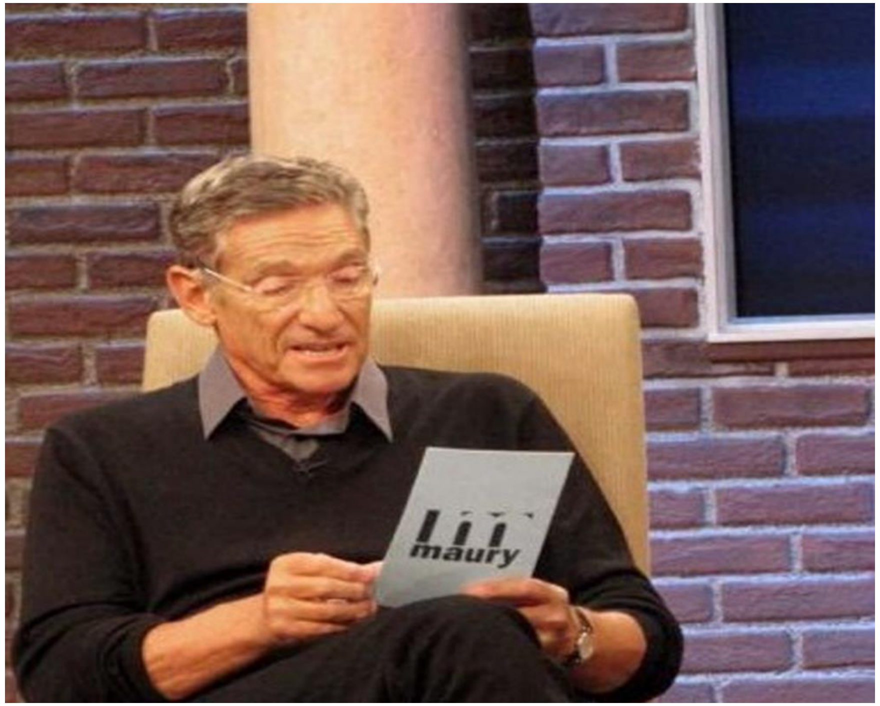
HW8 will be available soon. Due on 3/28 (Friday after returning from spring break)

Today’s draft problem

To be presented by today’s draftee.

Variation of Ex. 4d (with more accurate numbers)
An expert witness in a paternity suit testifies that the length (in days) of human gestation is approximately 280 days with a standard deviation of 14 days. The defendant in the suit (a long haul trucker) is able to prove that he was not in the mother’s time-zone for a period of 40 days that began 300 days before the child’s birth.

If the defendant were the father, what is the probability that the mother had the kind of exceptionally long or short gestation period implied by the evidence?



Recall - normal random variables

Definition (more-or-less from the book):

A continuous $\mathbb{R}$-valued random variable X (on a sample space S with probability measure P…) is called *normal* with parameters μ and σ^2 if its PDF is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

We showed last class that $\mathrm{E}[X] = \mu$ and $\mathrm{Var}(\mathrm{X}) = \sigma^2$.

The *standard normal* has parameters 0 and 1. Its CDF is denoted with Φ.

Table 5.1 Area $\Phi(x)$ Under the Standard Normal Curve to the Left of X.

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Normal approximation to binomial random variables

DeMoivre-Laplace limit theorem

If X_n is a binomial random variable with parameters (n, p), then for any $a < b$

$$\lim_{n\to\infty} P\left\{a \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b\right\} = \Phi(b) - \Phi(a)$$

Recall: $E[X_n] = np$ and $SD(X) = \sqrt{Var(X)} = \sqrt{np(1-p)}$

We won't prove this theorem. It's a special case of the Central Limit Theorem.

Monday’s draft problem

To be presented by Monday’s draftee (after return from spring break)

<u>Variation of Ex. 4h</u>

The 10 year average for Purdue’s undergraduate yield rate (that is, the fraction of admitted undergrad applicants who accept their admissions offer and show up) is 25.1%. Assuming that the admissions office’s goal was to have 9,815 freshman show up in fall 2024, they made 39,106 acceptances.

Use a normal approximation to determine the probability that 1,600 or more students than the goal show up.