STOR 435.001 Lecture 11

Continuous Random Variables - II

Jan Hannig

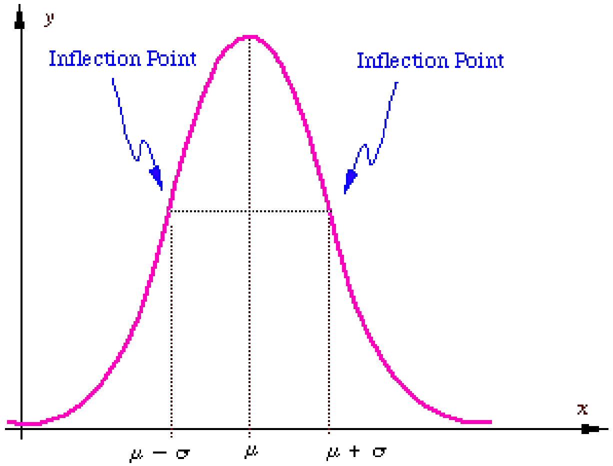
UNC Chapel Hill

Normal random variables

Definition

X is a normal random variable with parameters μ and σ^2 if its density is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$



Notation: $X = N(\mu, \sigma^2)$.

Note 1:

Importance: many real life distributions follow a normal density curve. We will see later in Section 8 that normal distributions arise in an important result known as Central Limit Theorem.

Examples

1. Standardized test scores: `http://nextsteptestprep.com/2012/04/13/average-lsat-score/`
2. Distribution of Physiological measurements: e.g. Blood pressure, beta carotene levels etc.
3. Used in ranking individuals in large companies. **Stack ranking**.
4. As a major driving component in (simplistic) financial models. **Black-Scholes model**.
5. Approximations to a wide array of other distributions. The Central Limit theorem. See end of these slides for a specific example.

Note 2:

$\frac{1}{\sqrt{2\pi\sigma^2}}$ is to have $\int_{-\infty}^{\infty} f(x)dx = 1$.

☼

Note 3:

If $X \sim N(\mu, \sigma^2)$ and a, b are two numbers, then the random variable $aX + b \sim N(a\mu + b, a^2\sigma^2)$.

☼

Terminology: $N(0,1)$ is called a *standard* normal random variable. Note that, if $X \sim N(\mu, \sigma^2)$, then

$$\frac{X-\mu}{\sigma} \sim N(0,1)$$

is a standard normal. This procedure is called standardizing, and allows all computations involving $N(\mu, \sigma^2)$ to be reduced to those for $N(0,1)$.

Normal random variables

Note 4:

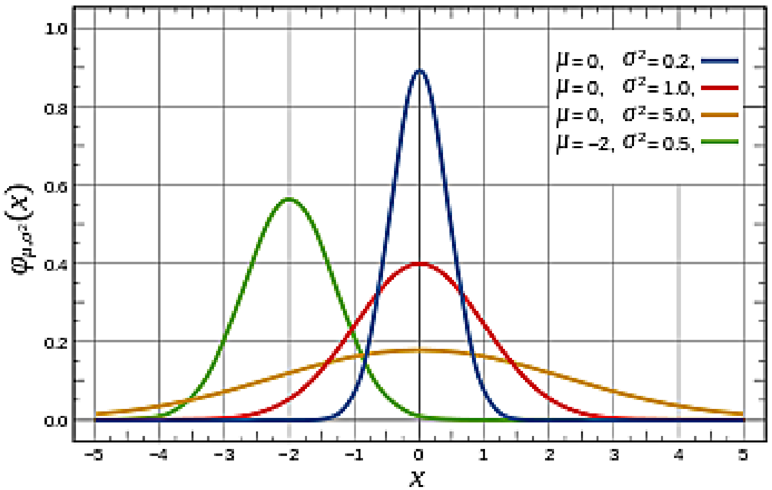
If $X \sim N(\mu, \sigma^2)$ then $\mu = EX$, $\sigma^2 = Var(X)$.

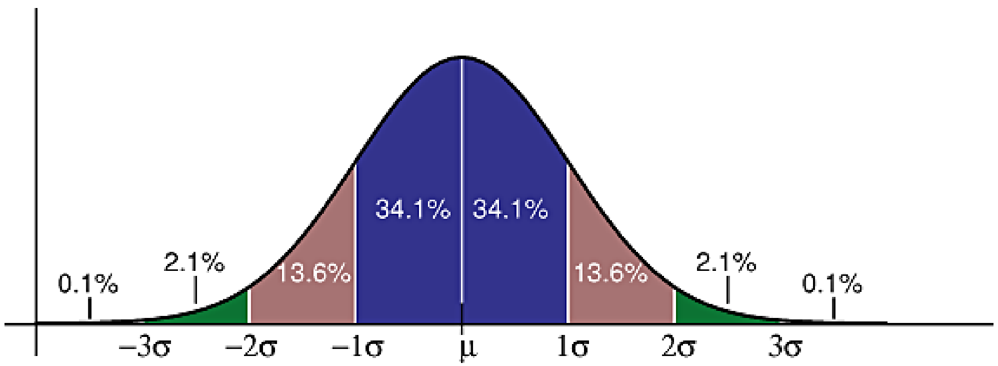
☼

Derivation of moments continued

☼

Normal random variables



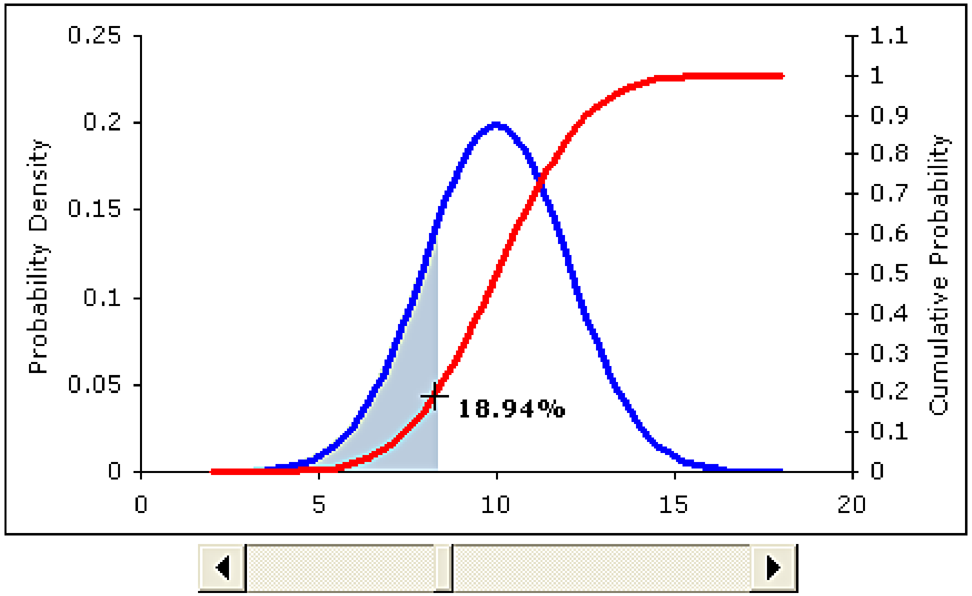


Note 5:

The cumulative distribution function of $N(0,1)$ is

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}}\, dy, \quad -\infty < x < \infty,$$

and does not have closed form formula.



For $x > 0$, $\Phi(x)$ are given in Table 5.1 of the textbook. For $x < 0$, use $\Phi(x) = 1 - \Phi(-x)$.

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

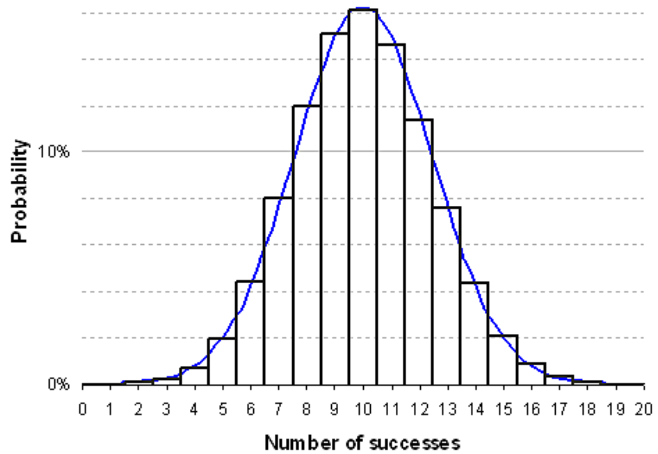
X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.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

Example 4b: If X is a normal random variable with parameters $\mu = 3$ and $\sigma^2 = 9$, find (a) $P(2 < X < 5)$; (b) $P(X > 0)$; (c) $P(|X - 3| > 6)$.

Normal approximations to binomial

Let $S_n \sim B(n,p)$ be the number of successes in n independent Bernoulli trials. Then, $ES_n = np$, $Var(S_n) = np(1-p)$ and, for large n,[1]

$$\frac{S_n - np}{\sqrt{np(1-p)}} \approx N(0,1).$$



[1] The approximation is quite good for $np(1-p) \geq 10$. Also, compare to Poisson approximations: p is not small here.

Continuity correction

Slightly subtle point: Binomial distribution is a **discrete distribution** while Normal is a continuous distribution. When approximating probabilities of the Binomial, need the *continuity correction*

☼

Schlitz vs Michelbob: Live taste test

See the following advertisement

https://adland.tv/commercials/schlitz-schlitz-vs-michelob-live-1981-060-usa

More reading

- http://wwnorton.tumblr.com/post/40864025234/the-naked-statistics-of-the-1981-schlitz-super-bowl-blin
- http://www.nytimes.com/1981/01/14/business/advertising-schlitz-to-challenge-michelob-s-taste.html

Example

Jan is the head of a major PR firm and plans an advertising campaign for a **VP** beer company. Leveraging on the fact that for non-specialty beers, most people cannot tell the difference between different brands of beer, he picks 100 affirmed drinkers of **Bond** beer and then does a live test on television where these individuals are given VP and Bond beer in unmarked mugs. Let X be the number of individuals who prefer VP beer. Find

(a) $P(X > 40)$

(b) $P(40 \leq X \leq 50)$.

☼

Solution contd

☼

Normal Approximation to the Binomial

Example 1: The ideal size of an incoming class of an awesome university is 4200. The college, knowing from past experience that, on the average, only .43 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 9500 students. Compute the probability that more than ($>$) 4200 first-year students attend this college.[2]

[2] I used the data from the Fall 2015 admissions profile of UNC.