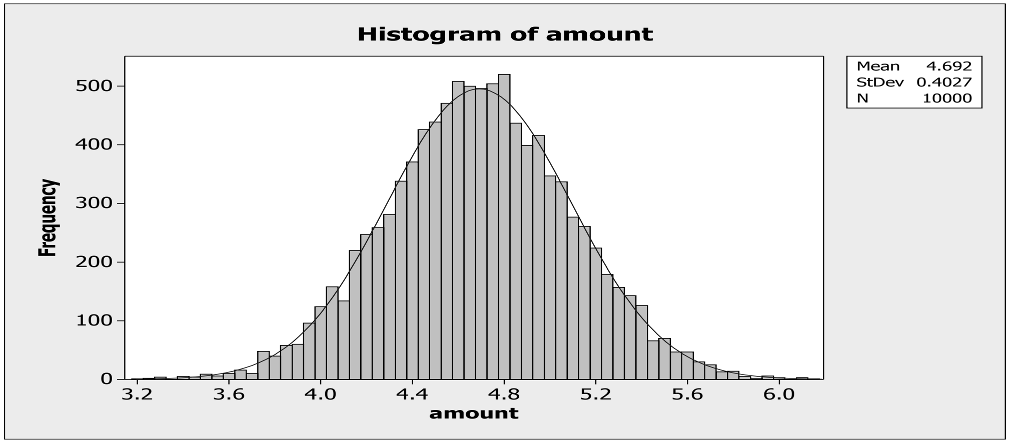
Stat 213: Intro to Statistics 8

H. Kim

Fall 2007

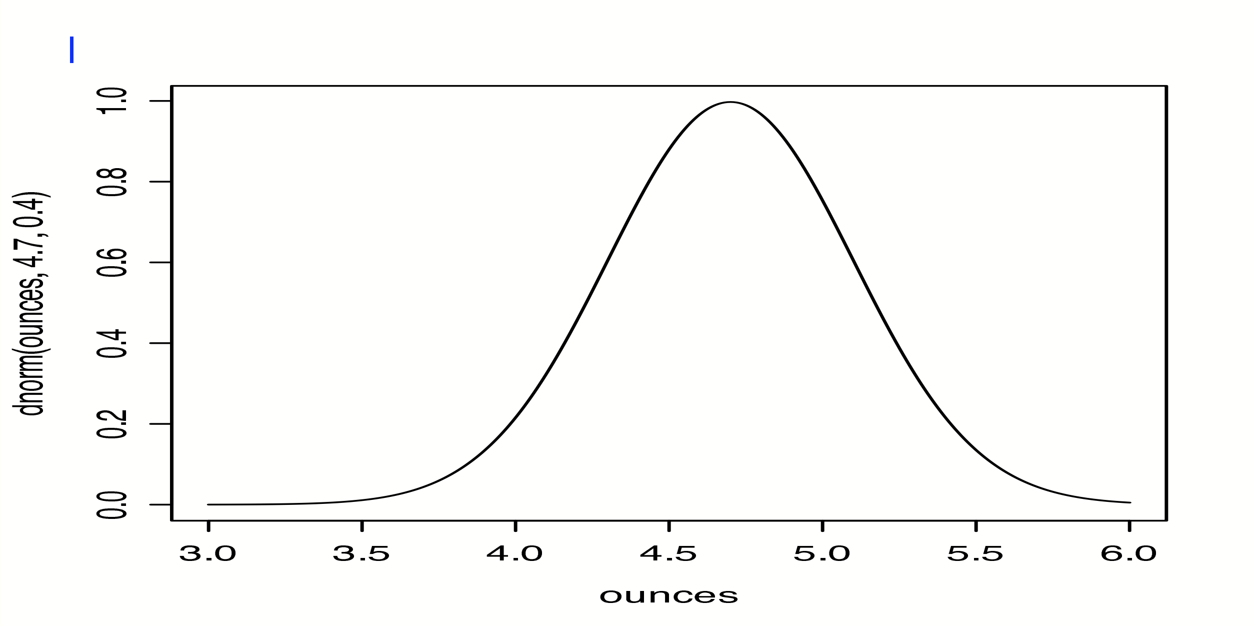
example 1

- A machine used to extract juice from oranges obtains an amount (ounces) from each orange



example 1

- A machine used to extract juice from oranges obtains an amount from each orange that is from a distribution, with mean of 4.70 ounces and a standard deviation of 0.40 ounces.



  - What is the probability that a randomly selected orange will contain between 4.70 and 5.00 ounces ?

the Normal probability distribution: $\mathcal{N}(\mu, \sigma^2)$

- Probability density function:
  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \; -\infty < x < \infty$
  - the **area** under a continuous probability distribution $= 1$
  - the probability that X will fall into a particular interval is equal to the **area** under the curve on the interval
  - $P(X = a) = 0$ for continuous random variables
  - $P(X \geq a) = P(X > a)$ or $P(X \leq a) = P(X < a)$

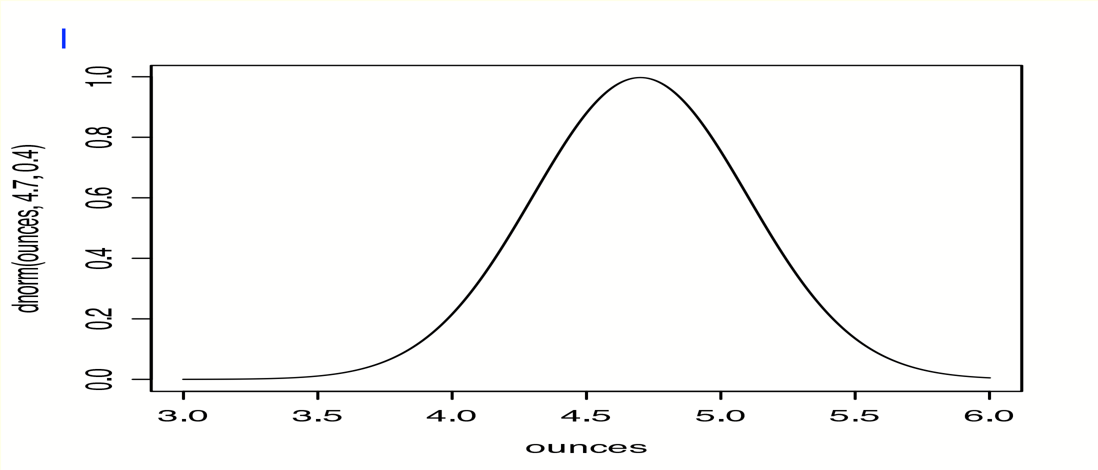
the Normal probability distribution: $\mathcal{N}(\mu, \sigma^2)$

- Probability density function:
  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \; -\infty < x < \infty$

  - cumulative probability density function: $F(x) = P(X \leq x)$: the area less than equal to x under the curve
  - $E(X) = \mu_X = \mu$
  - $V(X) = \sigma_X^2 = \sigma^2$

example 1: $\mathcal{N}(4.70, 0.4^2)$

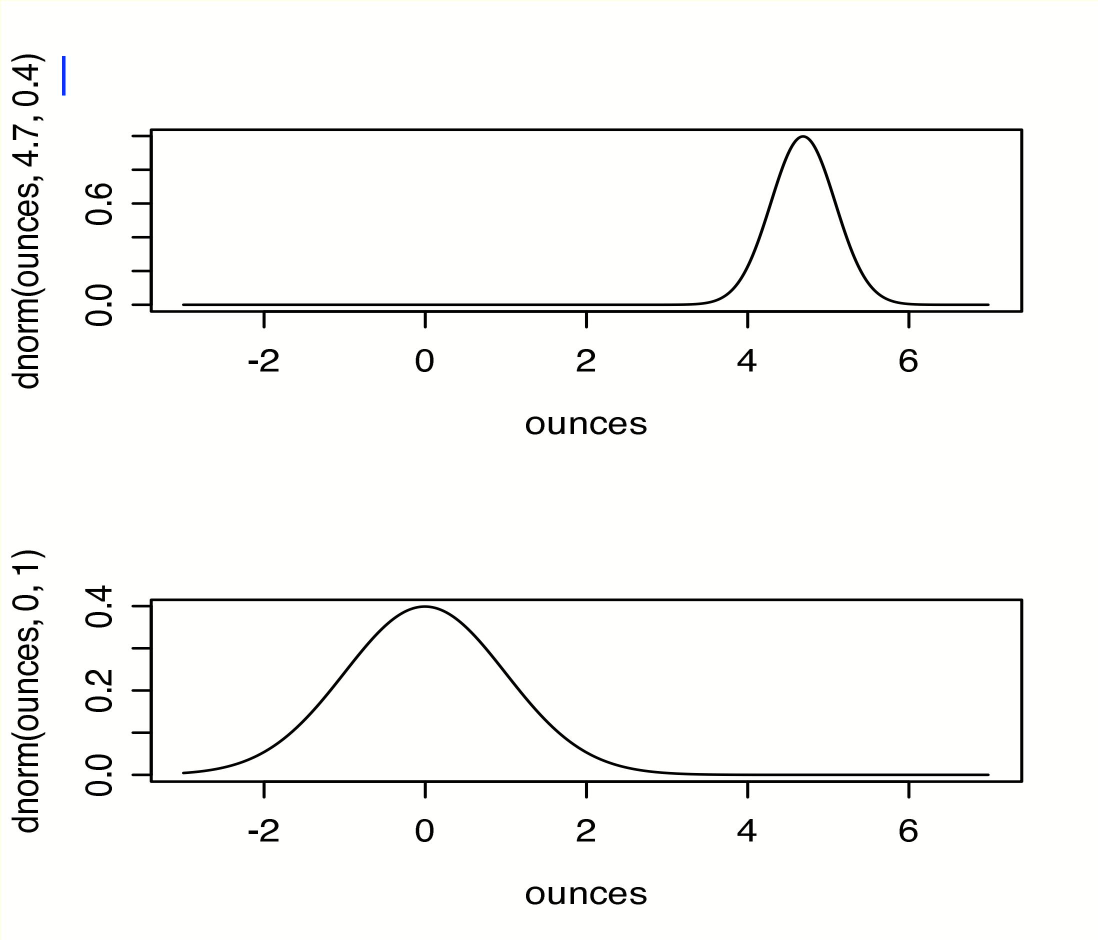
- Let X be the amount from each orange, that is approximately normally distributed, with mean of 4.70 ounces and a standard deviation of 0.40 ounces.
  - What is the probability that a randomly selected orange will contain between 4.70 and 5.00 ounces ?



  - $P(4.70 \leq X \leq 5.00) = P(X \leq 5.00) - P(X \leq 4.70)$

standardizing a normally distributed random variable

$$Z = \frac{X - \mu_X}{\sigma_X}$$

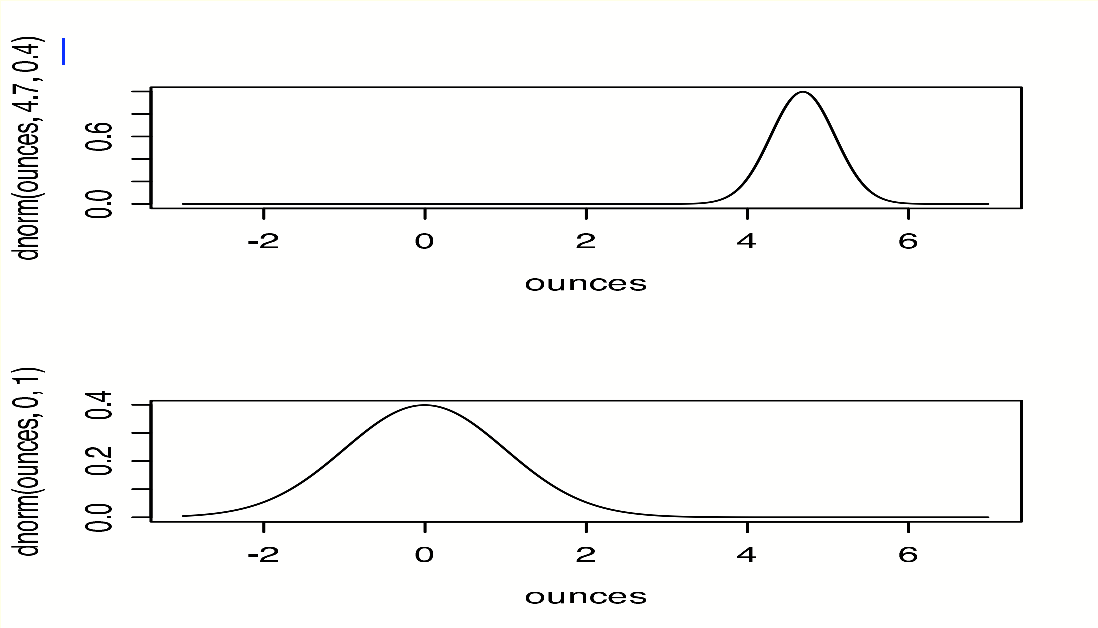


example 1: $\mathcal{N}(4.7, 0.4^2) \Longrightarrow \mathcal{N}(0,1)$

– What is the probability that a randomly selected orange will contain between 4.70 and 5.00 ounces ?

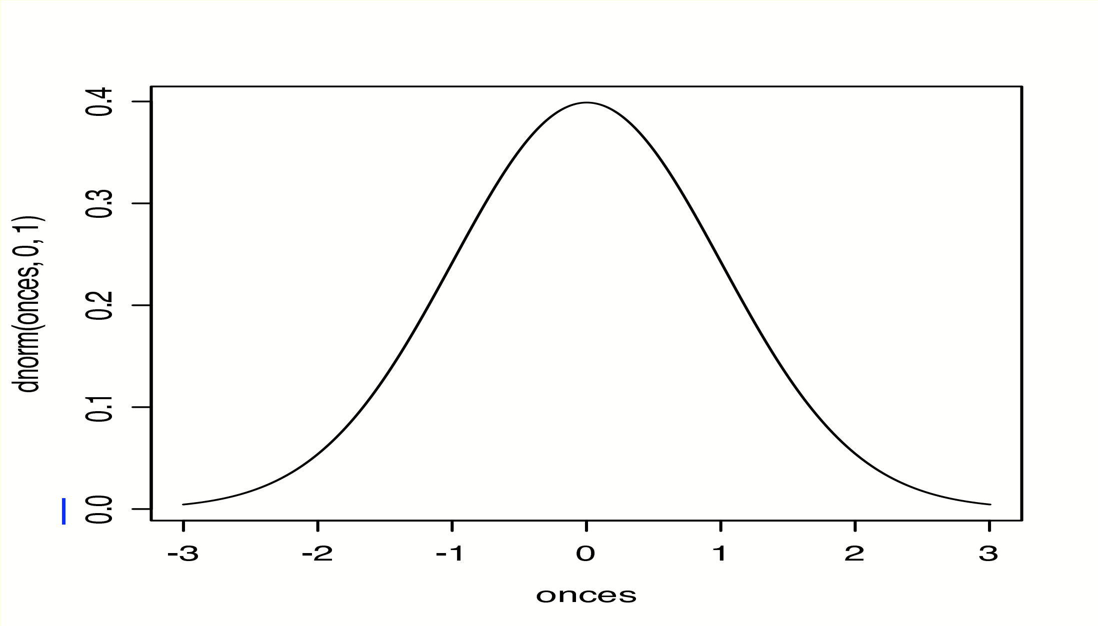
– z-score for 4.70: $z = \frac{4.70-4.70}{0.40} = 0.00$

– z-score for 5.00: $z = \frac{5.00-4.70}{0.40} = 0.75$



the standard Normal probability distribution: $Z \sim \mathcal{N}(0,1)$

- Probability density function: $f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \; -\infty < x < \infty$

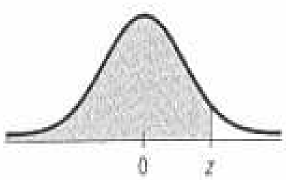


  - We will need a table to find the area.

table for $Z \sim \mathcal{N}(0,1)$

Appendix G • *Tables* **A-79**

Table Z (cont.)
Areas under the standard normal curve



	Second decimal place in z									
z	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

- $P(4.70 \leq X \leq 5.00) \Longrightarrow P(0 \leq Z \leq 0.75)$

  $= P(Z \leq 0.75) - P(Z \leq 0.00) = 0.7734 - 0.5 = 0.2734$

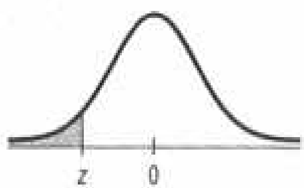
$$Z \sim \mathcal{N}(0,1) \Longrightarrow X \sim \mathcal{N}(\mu, \sigma^2)$$

- If Z has a standard normal distribution, i.e. $Z \sim \mathcal{N}(0,1)$, the the random variable X defined as $X = \mu + \sigma Z$ for some number of μ and $\sigma > 0$ has normally distributed with mean μ and variance σ^2, i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$.

table for $Z \sim \mathcal{N}(0,1)$

Table Z

Areas under the standard normal curve



Second decimal place in z										
0.09	*0.08*	*0.07*	*0.06*	*0.05*	*0.04*	*0.03*	*0.02*	*0.01*	*0.00*	z
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	−3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	−3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	−3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	−3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	−3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	−3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	−3.0

- $P(X \leq -3.4) = 0.0012$

example 2

- Find $P(-1 \leq Z \leq 1)$.

- Find $P(-0.5 \leq Z \leq 1)$.

- Find the value of z such that 0.95 of the area is within $\pm z_0$ standard deviation of the mean, i.e., find z_0 such that $P(-z_0 \leq Z \leq z_0) = 0.95$.

example 3

- If $P(Z \leq z_0) = 0.1190$, what is the z_0 ?

- If $X \sim \mathcal{N}(10, 4)$, find $P(X \geq 12.6)$.

- Let $X \sim \mathcal{N}(\mu, 1.53^2)$. If $P(X \geq 6.9944) = 0.07$, find the mean(μ).

example 4

- Let X be a normal random variable with mean μ and standard deviation σ . Find the probability that

  a. X is within one standard deviation of the mean

  b. X is within two standard deviation of the mean

example 5

- In one study, the length of time required to compute a test was normally distributed and the average was found to be 70 minute with a standard deviation of 12 minute. When should the test be terminated if we wish to allow sufficient time for 90% of the students to complete the test ?

example 6

The lifetimes of a particular type of light-bulb are approximately normally distributed with a mean of 1200 hours and a standard deviation of 140 hours.

a. Find the probability that a randomly chosen bulb will burn longer than 1400 hours.

b. Find the probability that a randomly chosen bulb will last between 1000 and 1050 hours.

example 6

(The lifetimes of a particular type of light-bulb $\sim \mathcal{N}(1200, 140)$.)

c. Find the probability that a randomly chosen bulb will last no larger than 1060 hours.

d. Twenty percent of the bulbs have a lifetime of less than what value ?

example 6

(The lifetimes of a particular type of light-bulb $\sim \mathcal{N}(1200, 140)$.)

e. Ten percent of the bulbs have a lifetime of greater than what value ?

f. At what number of hours should the warranty lifetime be set so that only 2% of bulbs must be replaced under warranty?

`the` $\mathcal{N}(\mu, \sigma^2)$ `approximation to the` $Binomial(n, p)$

- $Binomial(n, p) \Longrightarrow$
  - calculations: computer, or tables
  - if $np < 7$ and n: large, `Poisson` (λ) , $\lambda = np$
  - if $np \not< 7$ and n: large, $\mathcal{N}(\mu, \sigma^2)$, $\mu = np$ and $\sigma^2 = np(1 - p)$

    this approximation is adequate as long as n is large and p is not too close to 0 or 1 OR if $np > 5$ and $n(1 - p) > 5$.

continuity correction

helps account for the fact that you are approximating a discrete random variable with a continuous random variable

- Find the value of n and p and calculate $\mu = np$ and $\sigma^2 = np(1-p)$.
- Write down the probability you need in terms of X and locate the approximate area on the curve
- Correct the value of x by ± 0.5 to include the entire block of the probability for that value: continuity correction
- Convert the x values to z values using $z = \frac{x \pm 0.5 - np}{\sqrt{np(1-q)}}$
- Use the table to calculate the approximate probability

examples: continuity correction

- if $n = 50$:
  - $P(25 < X \leq 30) \Longrightarrow X = 26, \cdots, 30 \approx P(25.5 \leq X \leq 30.5)$
  - $P(X \leq 27) \Longrightarrow X = 0, \cdots, 27 \approx P(X \leq 27.5)$
  - $P(X > 30) \Longrightarrow X = 31, \cdots, 50 \approx P(X \geq 30.5)$
  - $P(27 < X < 31) \Longrightarrow X = 28, \cdots, 30 \approx P(27.5 \leq X \leq 30.5)$

example 7

- Let $X \sim Binomial(100, 0.5)$. Find $P(X \geq 60)$.

example 8

- Let 21% of North American deaths are due to cancer. What is the probability that 9 or more of the next 50 unrelated deaths are due to cancer ?