

STT351 - 004 and 005.

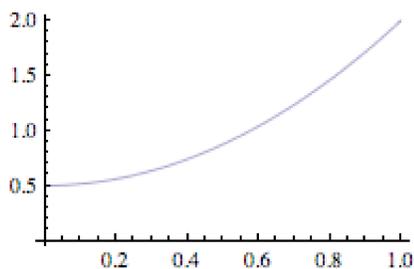
Week of Monday, September 29, 2008. The following hw is due Wednesday at the end of class.

Sec.4.1. Probability density $f(x)$.

a. Plot $f(x) := 0.5 + 1.5x^2$, $0 < x < 1$, 0 elsewhere.

```
f[x_] := 1/2 + 3 x^2 / 2
```

```
Plot[f[x], {x, 0, 1}, AxesOrigin -> {0, 0}]
```



b. Verify that f is nowhere negative and integrates to one (qualifies as a pr density).

```
Integrate[f[x], {x, 0, 1}]
```

1

c. Determine $P(X > 0.2)$ (integral of f over $x > 0.2$).

```
Integrate[f[x], {x, 0.2, 1}]
```

0.896

d. Determine $P(X^4 < 0.6)$ equal to $P(X < 0.60^{.25})$.

```
Integrate[f[x], {x, 0, .6^.25}]
```

0.780922

e. Determine $P(.4 < X < .7)$.

```
Integrate[f[x], {x, .4, .7}]
```

0.2895

f. Determine $P(3X + 2 < 3.5) = P(X < \frac{3.5-2}{3})$

```
Integrate[f[x], {x, 0, (3.5 - 2) / 3}]
```

0.3125

- g. Determine $P(|X - 0.2| < 0.3) = P(|X - 0.2| < 0.3)$
 $= P(0.2 - 0.3 < X < 0.2 + 0.3) = P(0 < X < 0.3)$ (since $X > 0$)

```
Integrate[f[x], {x, 0, 0.3}]
0.1635
```

- h. Determine $P(X = 0.3)$.

```
Integrate[f[x], {x, 0.3, 0.3}]
0. (1/2 + 3x^2/2)
```

A curious answer from *Mathematica*. It alerts us to trouble but would suggest 0 as the answer. The correct answer is 0 since the integral over a single point is zero.

Sec.4.2. Cumulative probability distribution $F(x)$, obtaining density $f(x)$ from F , percentiles and F , $E X$ from f , $\text{Var } X$ from f .

- i. Determine value c with $P(X < c) = 0.5$ (median of X).

```
F[x_] := Integrate[f[x], {x, 0, x}]
F[x]
x/2 + x^3/2
```

You must solve $F[x] = 0.5$. It is the solution of a cubic equation.

```
Solve[F[x] == 0.5, x]
{{x -> -0.341164 - 1.16154 i}, {x -> -0.341164 + 1.16154 i}, {x -> 0.682328}}
```

The only real valued solution is therefore median = 0.682328.

- j. Determine value c with $P(X < c) = 0.8$ (80 th percentile of X).

```
Solve[F[x] == 0.8, x]
{{x -> -0.445744 - 1.26335 i}, {x -> -0.445744 + 1.26335 i}, {x -> 0.891488}}
```

The only real root is 80th percentile = 0.891488.

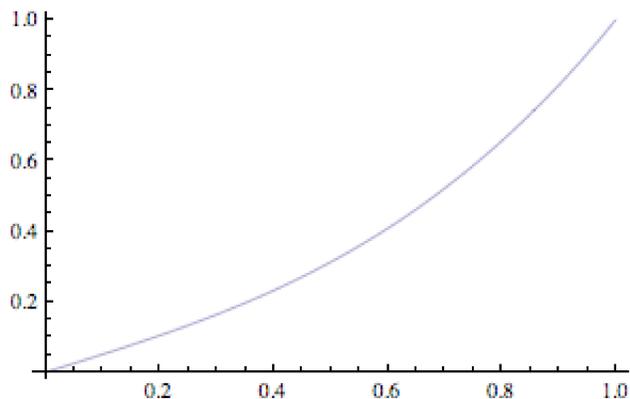
- k. Determine and plot $F(x)$. Notice that $F(0) = 0$ and $F(1) = 1$ which tells us that the probability distribution places all of its probability on $[0, 1]$.

```
F[x_] := Integrate[f[x], {x, 0, x}]
```

```
F[x]
```

$$\frac{x}{2} + \frac{x^3}{2}$$

```
Plot[F[x], {x, 0, 1}]
```



You should be able to obtain the median and 80th percentile by inspection of this plot.

1. Verify that $d/dx F(x)$ returns $f(x)$ (fundamental theorem of calculus). This is true on a set of x values having probability one (a nice clean interpretation offered by probability theory).

```
F[x]
```

$$\frac{x}{2} + \frac{x^3}{2}$$

```
F'[x]
```

$$\frac{1}{2} + \frac{3x^2}{2}$$

- m. Determine $E X$.

```
Integrate[x f[x], {x, 0, 1}]
```

$$\frac{5}{8}$$

If you wish to see the indefinite integral here is is (just put in the upper limit 1 for x to get $E X$).

```
Integrate[x f[x], {x, 0, x}]
```

$$\frac{x^2}{4} + \frac{3x^4}{8}$$

- n. Determine $E X^2$.

Integrate[$x^2 f[x]$, { x , 0, 1}]

$$\frac{7}{15}$$

o. Determine Var X from (m), (n).

Integrate[$x^2 f[x]$, { x , 0, 1}]

$$\frac{7}{15}$$

$$7 / 15 - (5 / 8)^2$$

$$\frac{73}{960}$$

N[%]

$$0.0760417$$

p. Determine sd X from (o).

Sqrt[$7 / 15 - (5 / 8)^2$]

$$\frac{\sqrt{\frac{73}{15}}}{8}$$

N[%]

$$0.275757$$

q. Determine the cumulative distribution function G for random variable $Y = X^2$.

$$G(y) = P(Y < y) = P(X^2 < y) = P(-\sqrt{y} < X < \sqrt{y}) = P(0 < X < \sqrt{y}) \text{ since } X > 0.$$

G[y_1] := **F**[\sqrt{y}]

G[y]

$$\frac{\sqrt{y}}{2} + \frac{y^{3/2}}{2}$$

r. From (q), determine the probability density $g(y)$ for r.v. Y (differentiate G).

G'[y]

$$\frac{1}{4\sqrt{y}} + \frac{3\sqrt{y}}{4}$$

s. From $g(y)$ (of part (r)) determine $E Y$ (i.e. $E X^2$ as determined from the distribution of $Y = X^2$).

$$G[y_] := F[\sqrt{y}]$$

$$G[y]$$

$$\frac{\sqrt{y}}{2} + \frac{y^{3/2}}{2}$$

$$g[y_] := G'[y]$$

$$g[y]$$

$$\frac{1}{4\sqrt{y}} + \frac{3\sqrt{y}}{4}$$

$$\text{Integrate}[y g[y], \{y, 0, 1\}]$$

$$\frac{7}{15}$$

t. Verify that your answers to (n) and (s) are indeed the same. They are both given as 7/15.

Sec.4.3. Normal distributions (all normals are alike in sd units from their mean). Suppose that in a particular human population IQ scores are approximately normally distributed with mean 100 and sd 15.

A. Determine the standard score (z - score) of IQ 121.

$$z = (121 - 100) / 15$$

$$\frac{7}{5}$$

$$N[\%]$$

$$1.4$$

B. Determine the fraction of persons having IQ < 121. It will be the same as the fraction of standard normal scores (z scores) below z-score (121 - 100)/15. This requires the table A.3 in the front of your textbook.

```
n[z_] := N[Integrate[Exp[-r^2/2]/Sqrt[2 Pi], {r, -Infinity, z}]]
```

```
n[0]
```

```
0.5
```

```
n[1]
```

```
0.841345
```

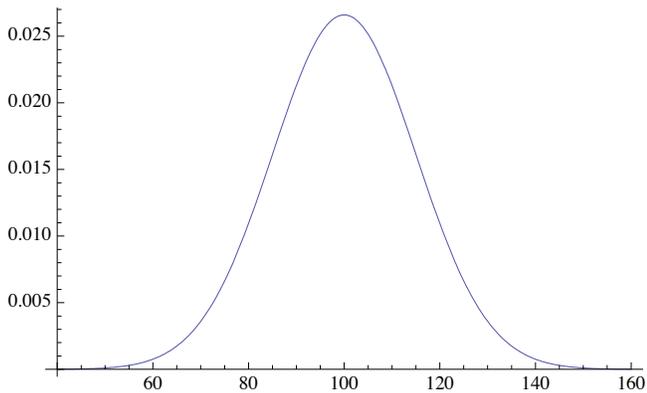
```
n[1.40]
```

```
0.919243
```

Entry $P(IW < 121) = 0.919243$ is (to table accuracy) found from $z = 1.40$ at row 1.4 and column .00 of the table in the front of your textbook.

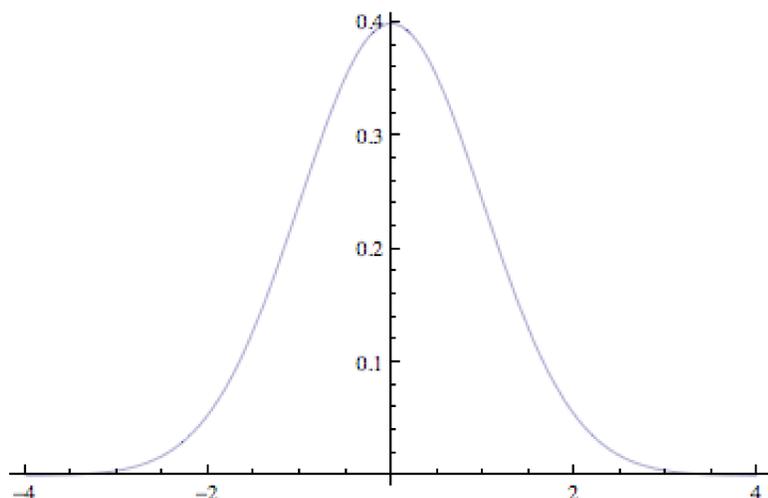
C. Sketch the IQ density with mean and sd labeled. Also sketch the standard normal density with mean and sd labeled. The two sketches may be in different scales.

```
Plot[Exp[-(x - 100)^2 / (2)] / (Sqrt[2 Pi]), {x, 100 - 60, 100 + 60}]
```



You must be able to identify (approximately, by eye) the mean 100 and sd 15 of IQ from the curve just above.

```
Plot[Exp[-r^2/2]/(sqrt(2 Pi)), {r, -4, 4}]
```



We've plotted each of the two normal curves out to 4 standard deviations. Since *Mathematica* tries to choose the scale for plotting (golden mean or whatever) the two plots actually appear identical except for printed scales.

D. Shade the region under the IQ density whose area equals $P(\text{IQ} < 121)$. Shade the area under the standard normal density whose area is also $P(\text{IQ} < 121)$, being sure to identify the standard score $z = 1.40$ (i.e. z-score) of 121.
(do so in the above curves)

E. The relationship between IQ and z is $z = (\text{IQ} - 100)/15$ which is exactly the same as $\text{IQ} = 100 + 15z$. In particular $\text{E IQ} = 100$ and $\text{Var IQ} = 225$ follow from the second relation. Show this is so.
(simple algebra since the second relation follows from the first by multiplying oth sides by 15 then adding 100)

F. To find the 90th percentile of Z from the z-table:

1. Enter 0.9 (or the closest inside table entry) to the z-table.

2. Read off z to two decimals.

```
Solve[n[z] == 0.90, z]
```

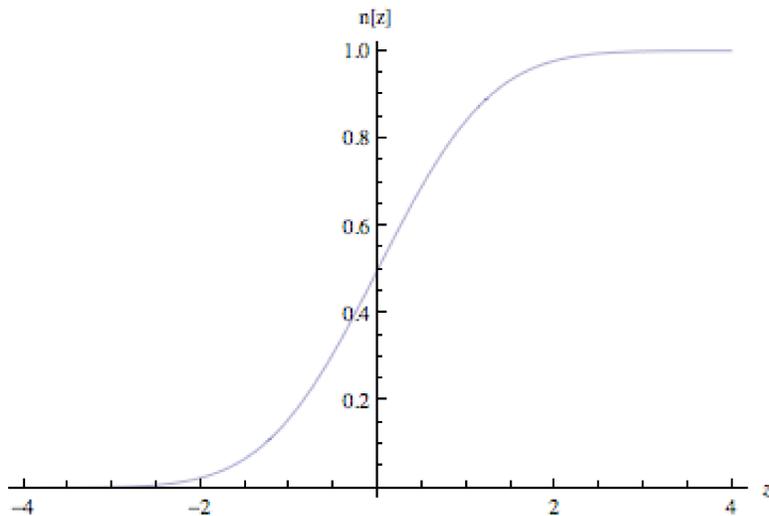
Solve::ifun : Inverse functions are being used by Solve, so some solutions may not be found; use Reduce for complete solution information. >>

```
{{z -> 1.28155}}
```

Confirm $z = 1.28$ from your talbe use.

You must also be able to obtain an approximation by ey simply by finding for which z the cumulative standard normal distribution function attains value 0.90. Check that it agrees with your table answer.

```
Plot[n[z], {z, -4, 4}, AxesLabel -> {z, "n[z]"}]
```



G. From (F) determine the 90th percentile of IQ.

3. From (F)(2) above calculate $\text{IQ} = 100 + 15 z$

$$100 + 15 \times 1.28155$$

$$119.223$$

H. Use the continuity correction to determine the 90th percentile of the binomial distribution with $n = 1000$ and $p = 0.3$. Instead I will use continuity correction to approximate $P(X = 301) \sim P(X \text{ in } (300.5, 301.5))$ (regarding X as approximately continuous).

$$n\left[\frac{(301.5 - 300)}{\sqrt{1000 \times 0.3 \times 0.7}}\right] - n\left[\frac{(300.5 - 300)}{\sqrt{1000 \times 0.3 \times 0.7}}\right]$$

$$0.0274587$$

$$\frac{(1000! / (301! (1000 - 301)!)) (0.3)^{301} (0.7)^{1000-301}}$$

$$0.0274296$$

The normal approximation (top) is close to the nearly exact $p(301)$ determined through direct evaluation of the binomial probability.

Sec.4.4. Exponential distribution (lifetime distribution).

I. For r.v. X having an exponential distribution and mean m , we have the useful relation $P(X > x) = e^{-x/m}$ for all values $x > 0$. Verify this by differentiation

$$\frac{d}{dx} P(X < x) = (1 - e^{-x/m})$$

to confirm the density (4.5) on page 157.

H[x_] := 1 - Exp[-x / mean]

H' [x]

$$\frac{e^{-\frac{x}{\text{mean}}}}{\text{mean}}$$

J. Use relation (I) to evaluate $P(X > 20)$ for mean $m = 10$.

Exp[-20 / 10]

$$\frac{1}{e^2}$$

N[%]

0.135335

K. What is (for any mean m) $P(X > m)$. That is, if lifetimes are exponentially distributed in a given population, what fraction of the population lives longer than the mean lifetime?

Exp[-mean / mean]

$$\frac{1}{e}$$

N[%]

0.367879

L. Again, assuming exponential lifetime distribution, if I am alive at time x (i.e. conditional on the event $X > x$) what is the conditional probability that I live an additional time r ? That is, determine

$$P(X > x + r \mid X > x) = P(X > x + r \text{ and } X > x) / P(X > x).$$

The trick is to see that $P(X > x + r \text{ and } X > x) = P(X > x + r)$.

Exp[-(x + r) / mean] / Exp[-x / mean]

$$e^{-\frac{x+r}{\text{mean}} + \frac{x}{\text{mean}}}$$

Simplify[%]

$$e^{-\frac{r}{\text{mean}}}$$

M. Show that your answer to (L) is equal to $P(X > r)$. That is, if I am alive at time x then I am “born again” in the sense that my additional lifetime is just as if I had just been born. How can it be? Think of “death by an out of the blue event” such as living until a coin finally comes up heads. If you are alive it means the coin has not yet come up heads. Your future life is like starting fresh since the future coin tosses are independent of past tosses.

(obvious)

Sec.4.6. Probability plot.

N. I will post the exam 1 grades next Monday together with a probability plot for them. Does it seem from the plot that the exam 1 grades follow a normal distribution? Be sure to state your reasoning as to why or why not.
(not yet done)