Computing Integrals

Bagels

We study a simple example of scientific computing.

- Assume that you run a Bagel&Juice cafè
- Each night you have to decide how many bagels you should order for the next day
- If you order too few - your customers will not be satisfied, and you will loose income
- If you order too many - it will be a waste of food and you loose money

Bagels

One strategy could be to order so many bagels that you

- in 95% of the days have enough
- in 5% of the days have to disappoint the last few customers

The question is therefore:

- How many bagels should be ordered to ensure that you have enough food 95% of the days?
- We shall not go into all the details of this statistical problem, but we will study the computational problem that arises in more depth

Bagels

- In statistics the number of sold bagels during one day is modeled with a probability density function - $f(x)$
- The probability, p, that the number of sold bagels is between a and b, is given by

$$p = \int_a^b f(x)dx \tag{1}$$

- The most common probability density function is the normal-distribution

$$f(x) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{(x-\bar{x})^2}{2s^2}}$$

where $\bar{x}$ is the mean value and s is called the standard deviation

Bagels

The mean value and the standard deviation can be estimated by the following procedure

- Assume that you have written down the number of sold bagels for a long period (n days), $x_1, x_2, x_3, \cdots, x_n$
- The mean value relative to this sample is given by

$$\bar{x} = \frac{1}{n} \sum_{j=1}^{n} x_j$$

- The sample standard deviation is given by

$$s = \sqrt{\frac{\sum_{j=1}^{n} (x_j - \bar{x})^2}{n-1}}$$

Bagels

Assume that these numbers for the Bagel& Juice cafè is $\bar{x} = 300$ and $s = 20$.

- The probability density function is then

$$f(x) = \frac{1}{\sqrt{2\pi}20} e^{-\frac{(x-300)^2}{2 \cdot 20^2}} \tag{2}$$

- The probability of selling less than b bagels is now given by

$$p = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}20} e^{-\frac{(x-300)^2}{2 \cdot 20^2}} dx \tag{3}$$

- How can p be computed for a given b?
- Note that $p \in [0, 1]$

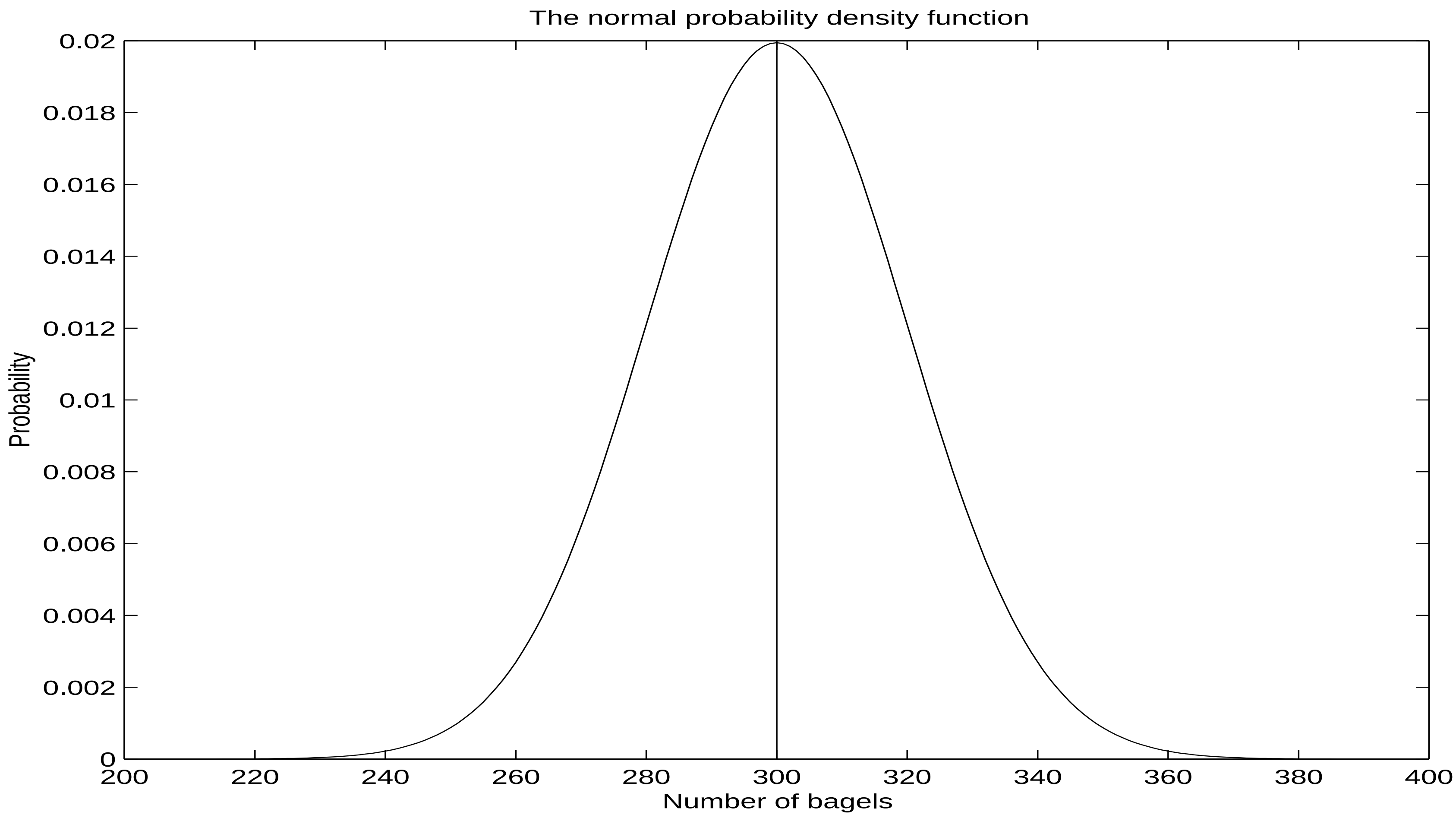

Figure 1: The figure illustrates the normal probability distribution in the case of $\bar{x} = 300$ and $s = 20$

Simplifications

- The integral can be divided into two parts

$$p = \int_{-\infty}^{300} f(x)dx + \int_{300}^{b} f(x)dx.$$

- Note that $f(x)$ is symmetric with respect to $x = 300$ in the sense that

$$f(300+x) = f(300-x)$$

- We conclude that

$$\int_{-\infty}^{300} f(x)dx = \int_{300}^{\infty} f(x)dx$$

Simplifications

- This means that

$$\int_{-\infty}^{300} f(x)dx = \frac{1}{2}\int_{-\infty}^{\infty} f(x)dx$$

- Every probability density function must fulfill

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

- which means that

$$\int_{-\infty}^{300} f(x)dx = \frac{1}{2}$$

Simplifications

- The computation of (3) can therefore be simplified to the computation of

$$p = \frac{1}{2} + \int_{300}^{b} f(x)dx \qquad (4)$$

- Note that the function $f(x) = \frac{1}{\sqrt{2\pi}20} e^{-\frac{(x-300)^2}{2\cdot 10^2}}$ is not analytically integrable (i.e. $\int_a^x f(\xi)d\xi$ can not be expressed by elementary functions)

Trapezoid method

- Generally we will study how to approximate definitive integrals of the form

$$\int_a^b f(x)dx$$

- Consider e.g. the function $f(x) = e^x$ and calculate

$$\int_1^2 e^x dx \tag{5}$$

- We will in the following pretend that this integral is not analytically integrable , and later use the exact analytical solution for comparison

Trapezoid method

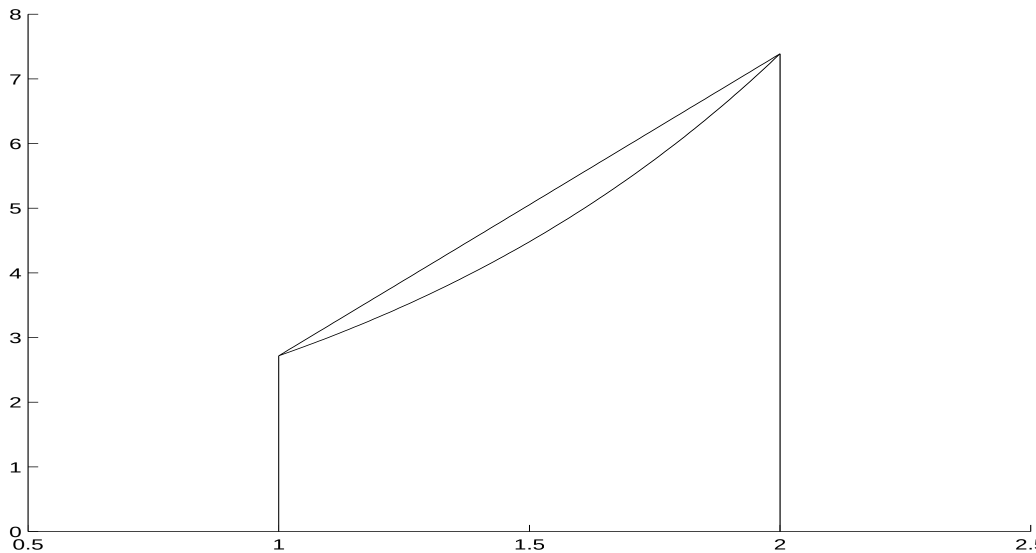

Figure 2: The figure illustrates how the integral of $f(x) = e^x$ (lower curve) may be approximated by a trapezoid on a given interval

Trapezoid method

- Let $y(x)$ be the straight line equal to f at the endpoints $x = 1$ and $x = 2$, i.e.

$$y(x) = e\left[1 + (e-1)(x-1)\right]$$

- Note that

$$y(1) = e = f(1)$$
$$y(2) = e^2 = f(2)$$

- Since $y(x) \approx f(x)$ we approximate the integral by

$$\int_1^2 e^x dx \approx \int_1^2 y(x) dx \tag{6}$$

Trapezoid method

We can now compute both integrals and compare the results

- Approximate

$$\int_1^2 y(x)dx = \int_1^2 e\left[1+(e-1)(x-1)\right]dx = \frac{1}{2}e + \frac{1}{2}e^2 \approx 5.0537$$

- Exact

$$\int_1^2 e^x dx = e(e-1) \approx 4.6708$$

Trapezoid method

The relative error is

- $$\frac{5.0537 - 4.6708}{5.0537} \cdot 100\% \approx 7.6\%$$

Trapezoid method

- Generally we can approximate the integral of f by

$$\int_a^b f(x)dx \approx \int_a^b y(x)dx \tag{7}$$

where $y(x)$ is a straight line equal to f at the endpoints, i.e.

$$y(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \tag{8}$$

- $y(x)$ is called the linear interpolation of f in the interval [a,b]

Trapezoid method

- Since y is linear, it is easy to compute the integral of this function

$$\int_a^b y(x)dx = \int_a^b \left[f(a) + \frac{f(b) - f(a)}{b-a}(x-a) \right] dx$$
$$= (b-a)\frac{1}{2}(f(a) + f(b))$$

- The trapezoid rule is therefore given by

$$\boxed{\int_a^b f(x)dx \approx (b-a)\frac{1}{2}(f(a) + f(b))} \qquad (9)$$

Example 1

- $f(x) = \sin(x), \;\; a = 1, \;\; b = 1.5$
- Trapezoid method

$$\int_1^{1.5} f(x)dx \approx (1.5-1)\frac{1}{2}(\sin(1)+\sin(1.5)) \approx 0.4597$$

- The exact value

$$\int_1^{1.5} f(x)dx = -\left[\cos(x)\right]_1^{1.5} = -(\cos(1.5)-\cos(1)) \approx 0.4696$$

- The relative error is

$$\frac{0.4696-0.4597}{0.4696} \cdot 100\% \approx 2.11\%$$

Trapezoid method

Now we approximate the integral using two trapezoids

- Choosing the middle point between a and b, $c = (a+b)/2$, we have that

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

- Using (9) on each integral gives

$$\int_a^b f(x)dx \approx \left[(c-a)\frac{1}{2}\left(f(a)+f(c)\right)\right] + \left[(b-c)\frac{1}{2}\left(f(c)+f(b)\right)\right]$$

Trapezoid method

- By using that

$$c-a=b-c=\frac{1}{2}(b-a),$$

we get

$$\boxed{\int_a^b f(x)dx \approx \frac{1}{4}(b-a)\left[f(a)+2f(c)+f(b)\right]} \qquad (10)$$

Example 2

- Using (10) on the problem considered in Example 1 gives

$$\int_1^{1.5} \sin(x)dx \approx \frac{1}{4} \cdot \frac{1}{2} [\sin(1) + 2\sin(1.25) + \sin(1.5)] \approx 0.4671$$

- The relative error of this approximation is

$$\frac{0.4696 - 0.4671}{0.4696} \cdot 100\% = 0.53\%$$

- This is significantly better than the approximation computed in in Example 1, where the error was 2.11%

Trapezoid method

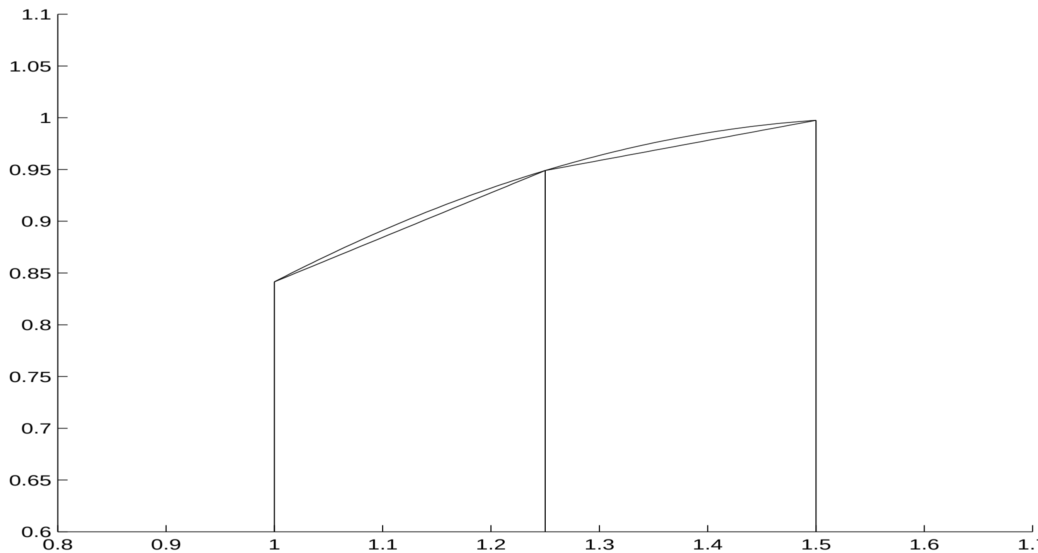

Figure 3: The figure illustrates how the integral of $f(x) = \sin(x)$ can be approximated by two trapezoids on a given interval

Trapezoid method

More generally we can approximate the integral using n trapezoids

- Let $h = \frac{b-a}{n}$
- Define $x_i = a + ih$
- The points

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

divide the interval from a to b into n subintervals of length h

Trapezoid method

- The integral has the following additive property

$$\begin{aligned}\int_a^b f(x)dx &= \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{n-1}}^{x_n} f(x)dx \\ &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx \qquad (11)\end{aligned}$$

- We use (9) on each integral, i.e.

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx (x_{i+1} - x_i)\frac{1}{2}\left[f(x_i) + f(x_{i+1})\right]$$

Trapezoid method

Since $h = x_{i+1} - x_i$, we get

$$\int_a^b f(x)dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx$$

$$\approx \sum_{i=0}^{n-1} \frac{h}{2}[f(x_i) + f(x_{i+1})]$$

$$= \frac{h}{2}\Big([f(x_0) + f(x_1)] + [f(x_1) + f(x_2)] + [f(x_2) + f(x_3)]$$

$$+ \cdots + [f(x_{n-2}) + f(x_{n-1})] + [f(x_{n-1}) + f(x_n)] \Big)$$

$$= h\Big[\frac{1}{2}f(x_0) + f(x_1) + f(x_2) + \cdots$$

$$\cdots + f(x_{n-2}) + f(x_{n-1}) + \frac{1}{2}f(x_n)\Big]$$

Trapezoid method

Written more compactly

$$\boxed{\int_a^b f(x)dx \approx h\left[\frac{1}{2}f(x_0) + \sum_{i=1}^{n-1} f(x_i) + \frac{1}{2}f(x_n)\right]} \qquad (12)$$

Example 3

The integral considered in Example 1 with $n = 100$.

- $h = \frac{b-a}{n} = \frac{0.5}{100} = 0.005$
- We get

$$\int_1^{1.5} \sin(x)dx \approx 0.005 \left[\frac{1}{2}\sin(1) + \sin(1.005) + \cdots + \frac{1}{2}\sin(1.5) \right]$$
$$= 0.469564$$

- The relative error is

$$\frac{0.469565 - 0.469564}{0.469565} \cdot 100\% = 0.0002\%$$

Example 4

Calculate $\int_0^1 f(x)dx$, where $f(x) = (1+x)e^x$

- The exact integral is

$$\int_0^1 (1+x)e^x dx = [xe^x]_0^1 = e$$

- Define $T_h = h\left[\frac{1}{2}f(0) + \sum_{i=1}^{n-1} f(x_i) + \frac{1}{2}f(1)\right]$
- where n is given and $h = \frac{1}{n}$ and $x_i = ih$ for $i = 1, \ldots, n$
- We want to study the error defined by

$$E_h = |e - T_h|$$

Example 4

n	h	E_h	E_h/h^2
1	1.0000	0.5000	0.5000
2	0.5000	0.1274	0.5096
4	0.2500	0.0320	0.5121
8	0.1250	0.0080	0.5127
16	0.0625	0.0020	0.5129
32	0.0313	0.0005	0.5129
64	0.0156	0.0001	0.5129

Table 1: The table shows the number of intervals, n, the length of the intervals, h, the error, E_h, and E_h/h^2

Example 4

- From the table it seems that

$$\frac{E_h}{h^2} \approx 0.5129$$

for small values of h

- That is

$$E_h \approx 0.5129h^2 \tag{13}$$

- This means that we can get as accurate approximation as we want

Example 4

- Assume that you want $E_h \leq 10^{-5}$
- then $0.5129h^2 \leq 10^{-5}$
- or $h \leq 0.0044$
- This means that $n = 1/h \geq 226.47$
- n has to be an integer, so therefore we set $n = 227$ to obtain the desired accuracy

Example 5

We want to test the trapezoid method for the following three integrals:

- $\int_0^1 x^4 dx$
- $\int_0^1 x^{20} dx$
- $\int_0^1 \sqrt{x} dx$
- Let E_h denote the error for a given value of h, i.e.

$$E_h = \left| \int_a^b f(x)dx - h\left[\frac{1}{2}f(x_0) + \sum_{i=1}^{n} f(x_i) + \frac{1}{2}f(x_n)\right]\right|,$$

where $h = \frac{b-a}{n}$ and $x_i = a + ih$ for $i = 0, \ldots, n$

Example 5

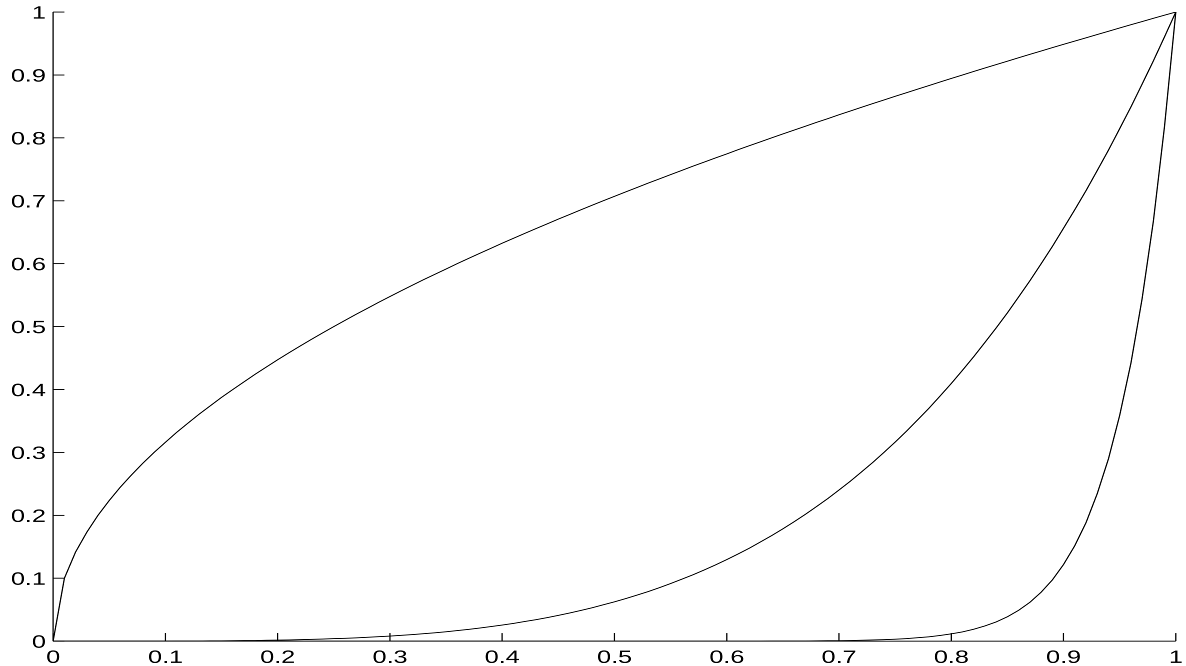

Figure 4: The figure shows the graph of $\sqrt{x}$ (upper), x^4 (middle) and x^{20} (lower)

Example 5

	$\int_0^1 x^4 dx = \frac{1}{5}$		$\int_0^1 x^{20} dx = \frac{1}{21}$		$\int_0^1 \sqrt{x} dx = \frac{2}{3}$	
h	$10^5 E_h$	E_h/h^2	$10^5 E_h$	E_h/h^2	$10^5 E_h$	E_h/h^2
0.01	3.33	0.33	16.66	1.67	20.37	2.04
0.005	0.83	0.33	4.17	1.67	7.25	2.90
0.0025	0.21	0.33	1.04	1.67	2.57	4.17
0.00125	0.05	0.33	0.26	1.67	0.91	5.84

Table 2: The table shows how accurate the trapezoidal method is for approximating three definite integrals.

Example 5

Conclusions

- In the two first integrals $\frac{E_h}{h^2}$ seems to be constant
- The constant is smaller for x^4 than for x^{20}
- The approximate integral of $\sqrt{x}$ on [0,1], seems to converge towards the correct value as $h \longrightarrow 0$, but $\frac{E_h}{h^2}$ increases with decreasing h

Trapezoid method

- We have studied several examples where the exact integral is obtainable
- In practice these examples are not so interesting
- Numerical integration is more interesting on examples where analytical integration is impossible
- Recall that the problem faced by the Bagel&Juice has no analytical solution

Bagels

- Recall that we wanted to compute

$$p = p(b) = \frac{1}{2} + \int_{300}^{b} \frac{1}{\sqrt{2\pi}20} e^{-\frac{(x-300)^2}{2\cdot 20^2}} dx,$$

  and find the smallest possible value of $b = b^*$ such that $p = p(b^*) \geq 0.95$

- Then b^* corresponds to the number of bagels you should buy to be at least 95% certain that you have enough bagels for the next day

Bagels

We have that

$$p(b+1) = \frac{1}{2} + \int_{300}^{b+1} f(x)dx$$
$$= \frac{1}{2} + \int_{300}^{b} f(x)dx + \int_{b}^{b+1} f(x)dx$$
$$= p(b) + \int_{b}^{b+1} f(x)dx.$$

- Remembering that $p(300) = \frac{1}{2}$, we can now produce $p(b)$ for any $b \geq 300$ by the following iterative procedure

$$p(b+1) = p(b) + \int_{b}^{b+1} f(x)dx, \qquad b = 300, 301, \ldots$$

- The integral $\int_{b}^{b+1} f(x)dx$ can be computed with the

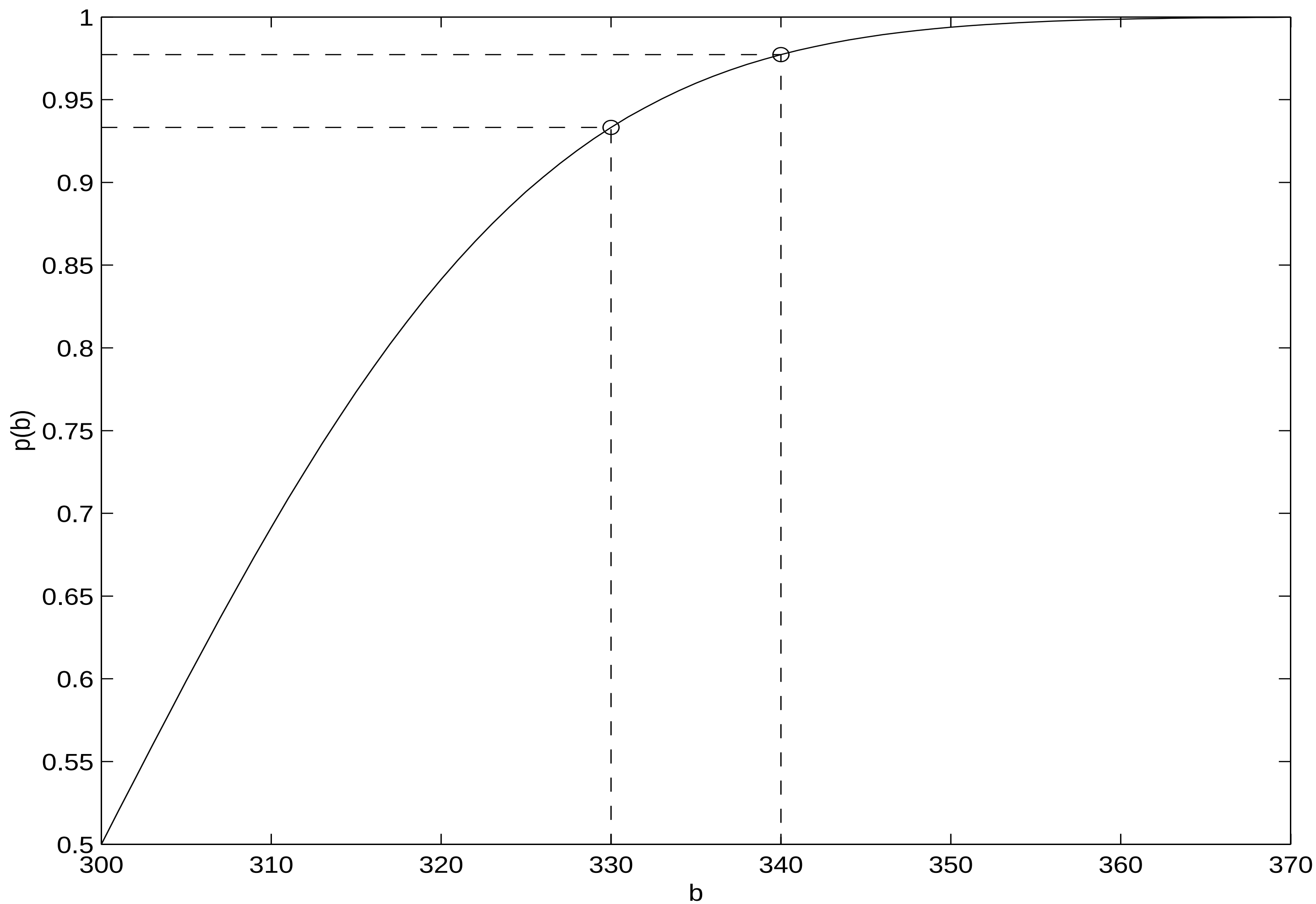

Figure 5: The graph of $p(b)$ where b is the number of bagels and $p(b)$ is the probability that the demand for bagels during a day is less than or equal to b.

Bagels

b	$p(b)$
331	0.939
332	0.945
333	0.951
334	0.955

Table 3: The table shows the the probability $p(b)$ for the demand of bagels at one particular day to be less than or equal to b.

Conclusion

- From Figure 5 we see that the desired value of b assuring sufficient supply in at least 95% of the days, is between 330 and 340
- Further more from Table 3 we observe that by choosing $b^* = 333$, we will get sufficient supplies at 95.1% of the days