EE126: Probability and Random Processes

Lecture 10: CDFs, Normal Random Variable

Abhay Parekh

UC Berkeley

February 17, 2011

Logistics

1. Midterm covers everything through today.
2. Use Piazzza more for clarifying conceptual issues.

Logistics

1. Midterm covers everything through today.
2. Use Piazzza more for clarifying conceptual issues.
3. Next week is Travel week...
    - Monday is President's Day so no office hours.
    - Abhay out for the week starting Wed.
    - Guy out all of next week.
    - I will have my regular office hour on Tuesday and Arash will have his on Friday.
    - Arash will teach on Thursday.

PDFs

Alice has three ways to get to a concert. With probability $1/3$ her friend, Bob will drive her there in time uniformly distributed over $[30, 45]$ minutes. With probability $2/3$ she will have to take the bus which takes time distributed with density $\alpha t - \beta$ for $t \in [50, 70]$. What are α, β and the density function for the time it will take Alice to get to the concert?

PDFs

Alice has three ways to get to a concert. With probability $1/3$ her friend, Bob will drive her there in time uniformly distributed over $[30, 45]$ minutes. With probability $2/3$ she will have to take the bus which takes time distributed with density $\alpha t - \beta$ for $t \in [50, 70]$. What are α, β and the density function for the time it will take Alice to get to the concert?

$$f_T(t) = \begin{cases} \frac{1}{45}, & t \in [30, 45]; \\ \frac{t}{300} - \frac{1}{6}, & t \in [50, 70]; \\ 0, & \text{o.w.} \end{cases}$$

Cumulative Distribution Functions

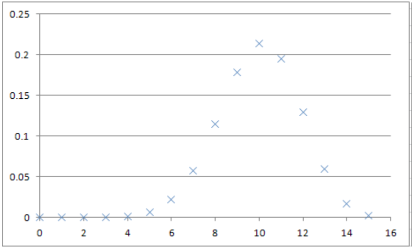
$$F_x(x) = Pr(X \leq x)$$

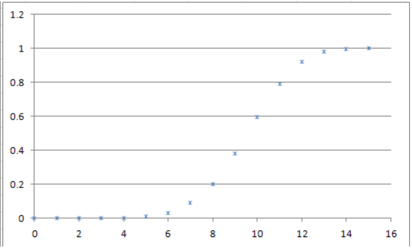
Works for continuous and discrete random variables.
The CDF is sometimes referred to as the **Probability Law** of the random variable.

Cumulative Distribution Functions

$$F_x(x) = Pr(X \leq x)$$

Works for continuous and discrete random variables.
The CDF is sometimes referred to as the **Probability Law** of the random variable. Discrete RV Example: Binomial: N=15, p=2/3:

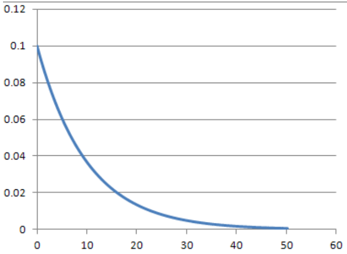


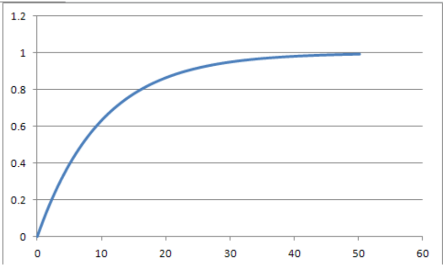


Cumulative Distribution Functions

$$F_x(x) = Pr(X \leq x)$$

Works for continuous and discrete random variables.
The CDF is sometimes referred to as the **Probability Law** of the random variable.
Continuous RV Example: Exponential $f_X(x) = \frac{1}{10}e^{-0.1x}$, $x \geq 0$:





Properties of the CDF

1. Non decreasing, tends to 1 as $x \to \infty$ and to 0 as $x \to -\infty$
2. If X is discrete:

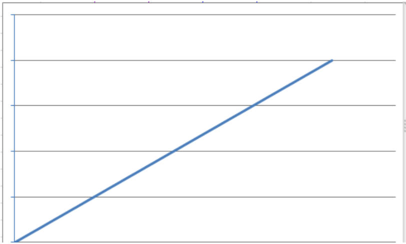
$$p_X(k) = F_X(k) - F_X(k-1)$$

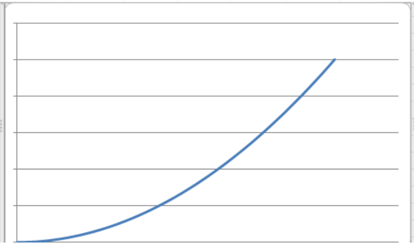
3. If X is continuous

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Example: Linear Distribution

$$f_X(x) = \begin{cases} \frac{2x}{a^2}, & x \in [0, a]; \\ 0, & \text{o.w.} \end{cases}$$





Exponential Random Variable

Since the geometric random variable is so useful, we want a non-negative memoryless continuous random variable:

$$P(X > t + s) = P(X > t)P(X > s)$$

Exponential Random Variable

Since the geometric random variable is so useful, we want a non-negative memoryless continuous random variable:

$$P(X > t + s) = P(X > t)P(X > s)$$

If $P(X > x) = e^{-\lambda x}$ for any $\lambda > 0$ then this is true. Then $F_X(x) = 1 - e^{-\lambda x}$ But

$$\begin{aligned} f_X(x) &= \frac{d(1 - e^{\lambda x})}{dx} \\ &= \lambda e^{-\lambda x} \end{aligned}$$

Are there more conditions on λ? We require that $\int_0^\infty f_X(x)dx = 1$:

Exponential Random Variable

Since the geometric random variable is so useful, we want a non-negative memoryless continuous random variable:

$$P(X > t + s) = P(X > t)P(X > s)$$

If $P(X > x) = e^{-\lambda x}$ for any $\lambda > 0$ then this is true. Then $F_X(x) = 1 - e^{-\lambda x}$ But

$$\begin{aligned} f_X(x) &= \frac{d(1 - e^{\lambda x})}{dx} \\ &= \lambda e^{-\lambda x} \end{aligned}$$

Are there more conditions on λ? We require that $\int_0^\infty f_X(x)dx = 1$:

$$\int_0^\infty \lambda e^{-\lambda x} dx = -|_0^\infty e^{-\lambda x} = 1$$

as long as $\lambda > 0$.

We call this the Exponential Distribution. Very Important!

Exponential Distribution

A non negative random variable X is exponentially distributed with parameter $\lambda > 0$ if

$$f_X(x) = \lambda e^{-\lambda x}$$

Exponential Distribution

A non negative random variable X is exponentially distributed with parameter $\lambda > 0$ if

$$f_X(x) = \lambda e^{-\lambda x}$$

- $F_X(x) = 1 - e^{-\lambda x}$
- $E[X] = \lambda \int_0^\infty x e^{-\lambda x} dx = \frac{1}{\lambda}$
- $var(X) = \frac{1}{\lambda^2}$.

Exponential Distribution

A non negative random variable X is exponentially distributed with parameter $\lambda > 0$ if

$$f_X(x) = \lambda e^{-\lambda x}$$

- $F_X(x) = 1 - e^{-\lambda x}$
- $E[X] = \lambda \int_0^\infty x e^{-\lambda x} dx = \frac{1}{\lambda}$
- $var(X) = \frac{1}{\lambda^2}$.

Note: The Laplace distribution is a kind of two-sided Exponential distribution. (Conditional on $x > 0$, it is exponential.)

Exponential and Geometric

CDF of Exponential r.v. X: is $1 - e^{-\lambda x}$
The CDF of a geometric rv, N: $\sum_{k=1}^{n} p(1-p)^{k-1} =$

Exponential and Geometric

CDF of Exponential r.v. X: is $1 - e^{-\lambda x}$
The CDF of a geometric rv, N: $\sum_{k=1}^{n} p(1-p)^{k-1} = 1 - (1-p)^n$.

Exponential and Geometric

CDF of Exponential r.v. X: is $1 - e^{-\lambda x}$
The CDF of a geometric rv, N: $\sum_{k=1}^{n} p(1-p)^{k-1} = 1 - (1-p)^n$.
Suppose we toss a coin every δ seconds with $p = 1 - e^{-\lambda\delta}$ for some $\delta << 1$. Then time to heads, X is geometric with

$$F_N(n) = 1 - e^{-n\lambda\delta}.$$

Exponential and Geometric

CDF of Exponential r.v. X: is $1 - e^{-\lambda x}$
The CDF of a geometric rv, N: $\sum_{k=1}^{n} p(1-p)^{k-1} = 1 - (1-p)^n$.
Suppose we toss a coin every δ seconds with $p = 1 - e^{-\lambda\delta}$ for some $\delta << 1$. Then time to heads, X is geometric with

$$F_N(n) = 1 - e^{-n\lambda\delta}.$$

But this means that

$$F_N(n) = F_X(n\delta)$$

For other values of x, the two distributions look very similar.

Example: Waiting at a post office

When you arrive there are two clerks both serving customers. You will be served when one these customers leaves. The amount of time a clerk spends with a customer is exponentially distributed with parameter λ. What is the probability that you are the last to leave the post office?

Example: Waiting at a post office

When you arrive there are two clerks both serving customers. You will be served when one these customers leaves. The amount of time a clerk spends with a customer is exponentially distributed with parameter λ. What is the probability that you are the last to leave the post office?
Suppose the first customer is done at time T_1, which is when you start being served. Since the exponential rv is memoryless,

Example: Waiting at a post office

When you arrive there are two clerks both serving customers. You will be served when one these customers leaves. The amount of time a clerk spends with a customer is exponentially distributed with parameter λ. What is the probability that you are the last to leave the post office?

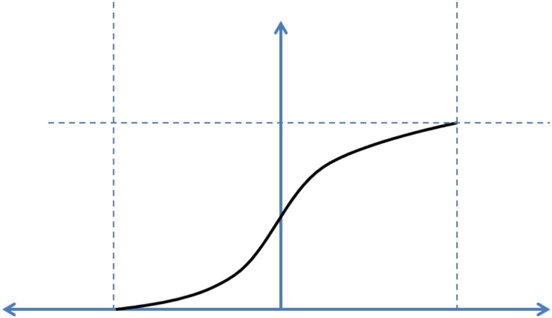
Suppose the first customer is done at time T_1, which is when you start being served. Since the exponential rv is memoryless, the remaining service time for both remaining customers (you and the other person) is the same. So prob you leave last is .5.

Example: Simulating any Random Variable from the Uniform Distribution

Suppose we can conduct an experiment that gives us values for X, distributed uniformly on $[0, 1]$, but we want to simulate another random variable Y, given by $F_Y(y)$.

Example: Simulating any Random Variable from the Uniform Distribution

Suppose we can conduct an experiment that gives us values for X, distributed uniformly on $[0, 1]$, but we want to simulate another random variable Y, given by $F_Y(y)$.



Derived Distributions

Suppose we know $F_X(x)$ and want to find $F_Y(x)$ where $Y = g(X)$.
Derive the CDF of Y from that of X.
Example: $F_X(x)$ is uniform on $[a, b]$, $a \geq 0$ and $Y = X^2$.

Derived Distributions

Suppose we know $F_X(x)$ and want to find $F_Y(x)$ where $Y = g(X)$. Derive the CDF of Y from that of X.

Example: $F_X(x)$ is uniform on $[a, b]$, $a \geq 0$ and $Y = X^2$.For $y \in [a^2, b^2]$:

$$P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F(-\sqrt{y})$$

$$F_Y(y) = \begin{cases} \frac{\sqrt{y}-a}{b-a}, & 0 \leq a \leq \sqrt{y} \leq b \\ 0 & \text{o.w.} \end{cases}$$

Derived Distributions

Suppose we know $F_X(x)$ and want to find $F_Y(x)$ where $Y = g(X)$. Derive the CDF of Y from that of X.

Example: $F_X(x)$ is uniform on $[a, b]$, $a \geq 0$ and $Y = X^2$.For $y \in [a^2, b^2]$:

$$P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F(-\sqrt{y})$$

$$F_Y(y) = \begin{cases} \frac{\sqrt{y}-a}{b-a}, & 0 \leq a \leq \sqrt{y} \leq b \\ 0 & \text{o.w.} \end{cases}$$

But what is $f_Y(y)$?

Derived Distributions

Suppose we know $F_X(x)$ and want to find $F_Y(x)$ where $Y = g(X)$.
Derive the CDF of Y from that of X.
Example: $F_X(x)$ is uniform on $[a, b]$, $a \geq 0$ and $Y = X^2$.For $y \in [a^2, b^2]$:

$$P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F(-\sqrt{y})$$

$$F_Y(y) = \begin{cases} \frac{\sqrt{y}-a}{b-a}, & 0 \leq a \leq \sqrt{y} \leq b \\ 0 & \text{o.w.} \end{cases}$$

But what is $f_Y(y)$?

$$f_Y(y) = \frac{dF_Y(ty)}{dy} = \frac{d(F_X(\sqrt{y}))}{dy} = \frac{1}{2(b-a)\sqrt{y}}$$

Note that

$$\int_{a^2}^{b^2} f_Y(y)dy = \Big|_{a^2}^{b^2} \frac{\sqrt{y}}{(b-a)}$$

Derived Distributions

Suppose $Y = aX + b$ and we know $F_X(x)$, $x \in \Re$. Then

$$Pr(Y \leq y) = Pr(aX + b \leq y) = \begin{cases} Pr(X \leq \frac{y-b}{a}), & \text{if } a > 0; \\ Pr(X \geq \frac{y-b}{a}), & \text{if } a < 0. \end{cases}$$

Derived Distributions

Suppose $Y = aX + b$ and we know $F_X(x)$, $x \in \Re$. Then

$$Pr(Y \leq y) = Pr(aX + b \leq y) = \begin{cases} Pr(X \leq \frac{y-b}{a}), & \text{if } a > 0; \\ Pr(X \geq \frac{y-b}{a}), & \text{if } a < 0. \end{cases}$$

For $a > 0$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(\frac{y-b}{a}) \\ &= \frac{1}{a} f_X(\frac{y-b}{a}) \end{aligned}$$

For $a < 0$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}(1 - F_X(\frac{y-b}{a})) \\ &= \frac{1}{-a} f_X(\frac{y-b}{a}) \end{aligned}$$

Derived Distribution for Linear Functions

If $Y = aX + b$, $a \neq 0$

$$f_Y(y) = \frac{1}{|a|} f_X(\frac{y-b}{a})$$

Derived Distribution for Linear Functions

If $Y = aX + b$, $a \neq 0$

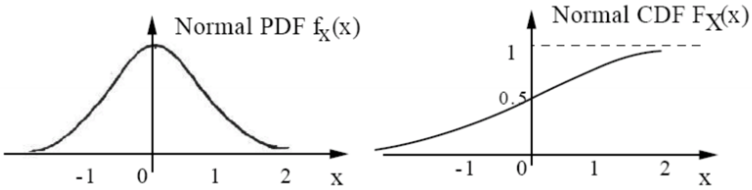
$$f_Y(y) = \frac{1}{|a|} f_X(\frac{y-b}{a})$$

Example: $X \sim$Exponential with parameter λ, $Y = aX + b$

$$f_y(y) = \frac{\lambda}{|a|} e^{-\frac{\lambda}{a}(y-b)}$$

Not an exponential rv, unless $b = 0$ and $a > 0$.

Standard Normal Variable



X is a standard normal variable if

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \quad x \in \Re$$

Since $f(x) = f(-x)$ and $\int_0^\infty x f_X(x) dx$ is finite, $E[X] = 0$.

Standard Normal Variable

We need to show that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx = 1$$

Standard Normal Variable

$$var(X) = E[X^2] = \frac{1}{\sqrt{2\pi}} \int_0^\infty x^2 e^{\frac{-x^2}{2}} dx$$

Integration by parts, $du = e^{\frac{-x^2}{2}}$, $v = x^2$

$$\begin{aligned} var(X) &= \frac{1}{\sqrt{2\pi}}(-xe^{\frac{-x^2}{2}})\Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx \\ &= 0 + 1 \end{aligned}$$

Bell Curve with zero mean and unit variance.

Standard Normal v/s Laplace

Recall Laplace $f_{\text{Laplace}}(x) = \frac{\lambda}{2} e^{-|x|\lambda}$ Laplace with $\lambda = \frac{1}{\sqrt{2}}$ has mean 0 and standard deviation 1. So,

$$f_{\text{std Laplace}}(x) = \frac{1}{2\sqrt{2}} e^{-\frac{|x|}{\sqrt{2}}}$$

$$f_{\text{std normal}}(x) = \frac{1}{2\sqrt{\pi}} e^{\frac{-x^2}{2}}$$

Compare density near the mean and in the tails...

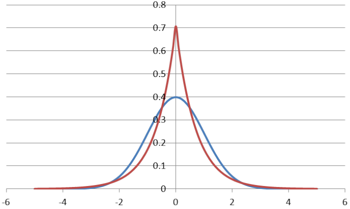
Standard Normal v/s Laplace

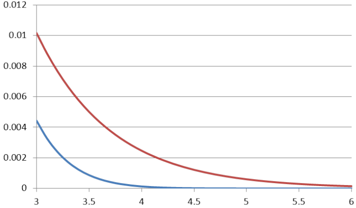
Recall Laplace $f_{\text{Laplace}}(x) = \frac{\lambda}{2} e^{-|x|\lambda}$ Laplace with $\lambda = \frac{1}{\sqrt{2}}$ has mean 0 and standard deviation 1. So,

$$f_{\text{std Laplace}}(x) = \frac{1}{2\sqrt{2}} e^{-\frac{|x|}{\sqrt{2}}}$$

$$f_{\text{std normal}}(x) = \frac{1}{2\sqrt{\pi}} e^{\frac{-x^2}{2}}$$

Compare density near the mean and in the tails...





Normal Random Variable

Bell curve with mean μ and variance σ^2.
Recall that for any RV X: $E(\frac{X-E[X]}{var(X)}) = 0$ and $var(\frac{X-E[X]}{\sqrt{var(X)}}) = 1$

Normal Random Variable

Bell curve with mean μ and variance σ^2.
Recall that for any RV X: $E(\frac{X-E[X]}{var(X)}) = 0$ and $var(\frac{X-E[X]}{\sqrt{var(X)}}) = 1$

So write $Y = \frac{X-\mu}{\sigma}$ and let Y be a standard normal variable. Then

$$P(Y \leq y) = P(\frac{X-\mu}{\sigma} \leq \frac{y-\mu}{\sigma}) = F_X(\frac{y-\mu}{\sigma})$$

$$f_x(y) = \frac{1}{\sigma} f_Y(\frac{y-\mu}{\sigma}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Normal Random Variable

Bell curve with mean μ and variance σ^2.
Recall that for any RV X: $E(\frac{X-E[X]}{var(X)}) = 0$ and $var(\frac{X-E[X]}{\sqrt{var(X)}}) = 1$

So write $Y = \frac{X-\mu}{\sigma}$ and let Y be a standard normal variable. Then

$$P(Y \leq y) = P(\frac{X-\mu}{\sigma} \leq \frac{y-\mu}{\sigma}) = F_X(\frac{y-\mu}{\sigma})$$

$$f_x(y) = \frac{1}{\sigma} f_Y(\frac{y-\mu}{\sigma}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Given any $\mu \in \Re$ and $\sigma^2 > 0$, the Normal/Gaussian R.V. X is given by:

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

Linear Function of a Gaussian

Suppose $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$. Then:

$$f_Y(y) = \frac{1}{|a|} f_X(\frac{y-b}{a})$$

Substituting:

$$f_Y(y) = \frac{1}{|a|\sqrt{2\pi\sigma^2}} e^{-\frac{(\frac{y-b}{a}-\mu)^2}{\sigma^2}}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{(y-b-a\mu)^2}{(a\sigma)^2}}$$

Linear Function of a Gaussian

Suppose $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$. Then:

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

Substituting:

$$f_Y(y) = \frac{1}{|a|\sqrt{2\pi\sigma^2}} e^{-\frac{(\frac{y-b}{a}-\mu)^2}{\sigma^2}}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{(y-b-a\mu)^2}{(a\sigma)^2}}$$

Thus $Y \sim N(b + a\mu, a^2\sigma^2)$.

Gaussian RV: Computing Probabilities

Example: Let the temperature of a city in the winter be X in Celsius. $X \sim N(2, 16)$. Find $P(-2 < X < 6)$.
Want to avoid computing the hairy integrals involved...Closed form of Normal CDF not available.
Procedure:

Gaussian RV: Computing Probabilities

Example: Let the temperature of a city in the winter be X in Celsius. $X \sim N(2, 16)$. Find $P(-2 < X < 6)$.
Want to avoid computing the hairy integrals involved...Closed form of Normal CDF not available.
Procedure:

1. Convert to a standard normal: $\hat{X} = \frac{X-2}{4}$
2. Express in terms of probability problem of $\hat{X}$:

$$P(-2 < X < 6) = P(\frac{-2-2}{4} < \hat{X} < \frac{6-2}{4}) = P(-1 < \hat{X} < 1)$$

Gaussian RV: Computing Probabilities

Example: Let the temperature of a city in the winter be X in Celsius. $X \sim N(2, 16)$. Find $P(-2 < X < 6)$.
Want to avoid computing the hairy integrals involved...Closed form of Normal CDF not available.
Procedure:

1. Convert to a standard normal: $\hat{X} = \frac{X-2}{4}$
2. Express in terms of probability problem of $\hat{X}$:

$$P(-2 < X < 6) = P(\frac{-2-2}{4} < \hat{X} < \frac{6-2}{4}) = P(-1 < \hat{X} < 1)$$

3. Look it up in a table of computed values of $F_{\hat{X}}(\hat{x})$
   We want to find $P(\hat{X} < 1) - P(\hat{X} < -1) = \phi(1) - \phi(-1)$. Tables typically only list positive values, so $\phi(-1) = 1 - \phi(1)$ by symmetry.

Gaussian RV: Computing Probabilities

Example: Let the temperature of a city in the winter be X in Celsius. $X \sim N(2,16)$. Find $P(-2 < X < 6)$.
Want to avoid computing the hairy integrals involved...Closed form of Normal CDF not available.
Procedure:

1. Convert to a standard normal: $\hat{X} = \frac{X-2}{4}$
2. Express in terms of probability problem of $\hat{X}$:

$$P(-2 < X < 6) = P(\frac{-2-2}{4} < \hat{X} < \frac{6-2}{4}) = P(-1 < \hat{X} < 1)$$

3. Look it up in a table of computed values of $F_{\hat{X}}(\hat{x})$
   We want to find $P(\hat{X} < 1) - P(\hat{X} < -1) = \phi(1) - \phi(-1)$. Tables typically only list positive values, so $\phi(-1) = 1 - \phi(1)$ by symmetry.
4. Answer: $2\phi(1) - 1 =$

Note: $2\phi(k) - 1$ tells you how likely it is that the outcome is within k standard deviations. $k = 2 : 0.9544$;k=3: 0.9974

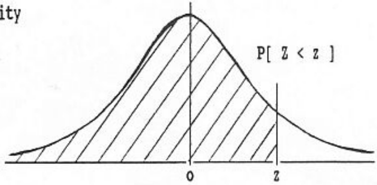
Table

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z i.e.

$$P[\, Z < z \,] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\, dZ$$



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	0.5159	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	0.7854
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
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830