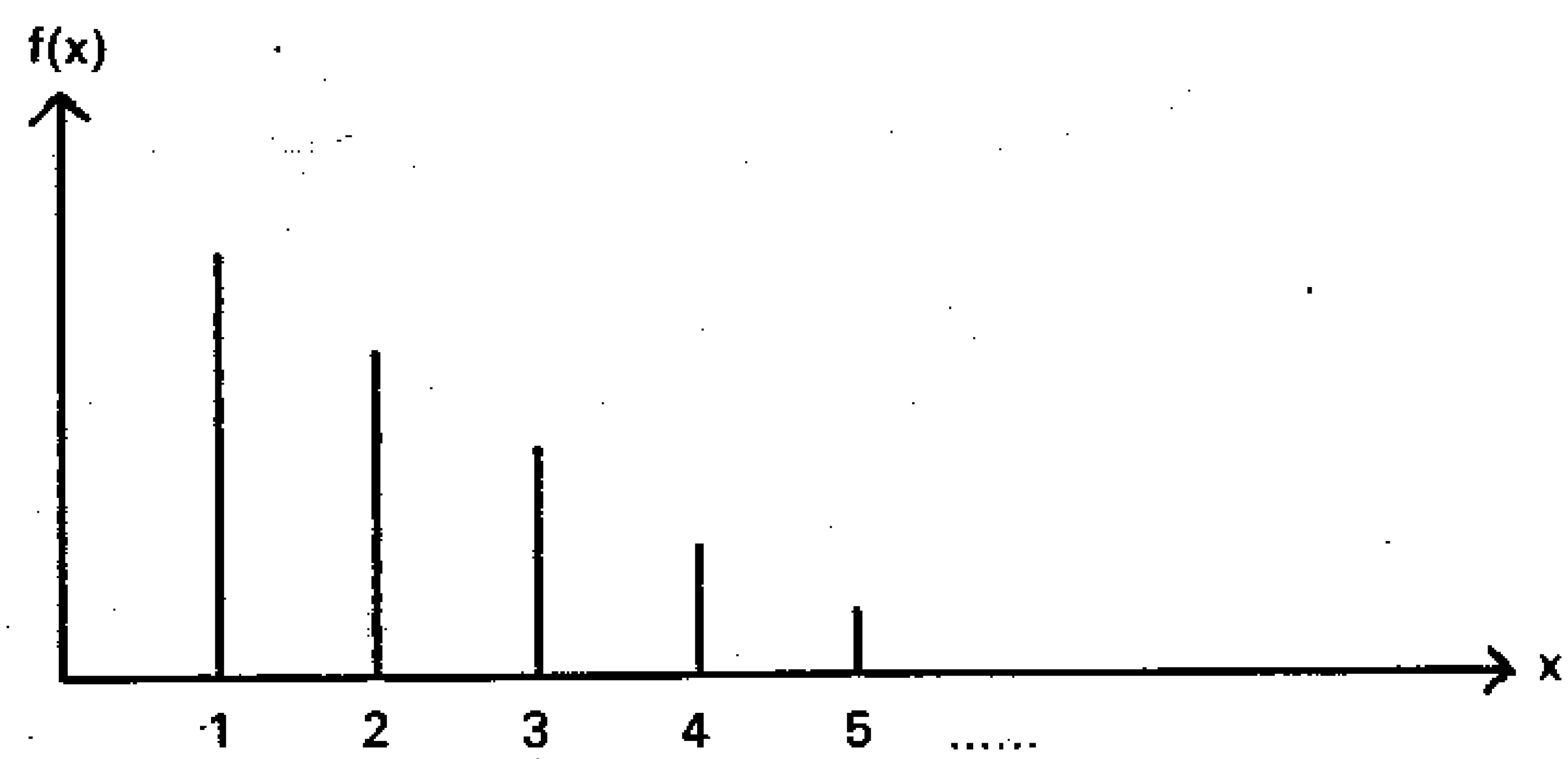
單元 20：隨機變量、離散及連續概率分佈

特定目標：

1. 能夠求出離散和連續概率分佈的期望值及方差。
2. 學習二項分佈和常態分佈及其日常應用。
3. 認識獨立常態變量綫性組合的性質。

內容	時間分配	教學建議
20.1 隨機變量 (*a*) 離散概率函數	6	教師可用簡單的例子介紹隨機變量，如擲毫(間斷隨機變量)及電燈泡的壽命(連續隨機變量)，使學生對此有一初步的認識。 教師可用擲兩枚硬幣的例子介紹離散概率函數 f(x) 為 $f(x)=P(X=x)$，其中 X 是一個離散隨機變量，x 是一個隨機變量的數值。 $f(x)=\begin{cases}0.25 & x=0\\ 0.5 & x=1\\ 0.25 & x=2\\ 0 & \text{其他}\end{cases}$ X(擲得正面的數目)是一個離散隨機變量，其數值可等於 0、1 或 2。 教師須強調以下的情況： $f(x)\geqslant 0$ 及 $\Sigma f(x)=1$ 同時亦要提醒學生，大楷 X 是隨機變量的記號，而小楷 x 是隨機變量的可能數值。 以下是另一離散隨機變量的例子。 *例* 擲一骰時，X 為投到六點的次數。 它的離散概率函數 f(x) 是 $f(x)=P(X=x)=\left(\frac{1}{6}\right)\left(\frac{1}{5}\right)^{x-1}$。明顯地，$\Sigma f(x)=\sum_{x=1}^{\infty}P(X=x)=\frac{\frac{1}{6}}{1-\left(\frac{5}{6}\right)}=1$

內容	時間分配	教學建議
		以綫條圖或組織圖來表示離散概率函數，更能幫助學生了解此概念。 
(*b*) 概率密度函數		在此階段，學生應已相當了解離散概率函數，教師可將此概念伸延至連續隨機變量，從而介紹連續概率密度函數 (p.d.f.) f(x)。學生須留意： $f(x)\geqslant 0$ 及 $\int_{-\infty}^{\infty}f(x)\,dx=1$ 學生須知道一個連續隨機變量 X 可取某一特定範圍內的任何數值，而它和 p.d.f. f(x) 的關係可用以下的公式表示： $P(a\leqslant X\leqslant b)=\int_a^b f(x)\,dx$ 事實上，p.d.f. 的圖表，可用統計學上連續數據的頻率曲綫來說明。

內容	時間分配	教學建議

此外，學生亦應注意：

$P(X=a)=0$ 及 $P(a\leqslant X<b)=P(a<X<b)=P(a<X\leqslant b)=P(a\leqslant X\leqslant b)$

以下是一些有關的例子：

例一(矩形分佈)
X的概率密度函數定義為

$$f(x)=\begin{cases} k & 0<x\leqslant 4 \quad k\text{是常數} \\ 0 & \text{其他} \end{cases}$$

從 $\int_{-\infty}^{\infty} f(x)\,dx=1$ 可找出k的數值。

同時 $P(-2<X\leqslant 1)=\int_{0}^{1} f(x)\,dx$

及 $P(X\geqslant 3)=1-\int_{0}^{3} f(x)\,dx=\int_{3}^{4} f(x)\,dx$

假設 $P(X\leqslant M)=b$，教師可要求學生以b表M。他們亦應注意到當 $b=0.5$ 時，M是中位數。

例二
某班機預期在早上八時正到達香港，但它實際到達的時間為早上 $(8+X)$ 時，X是一個隨機變量，其概率密度函數是：

$$f(x)=\begin{cases} \dfrac{3(4-x^2)}{32} & -2<x<2 \\ 0 & \text{其他} \end{cases}$$

教師可要求學生找出這班機在早上七時至八時，及在早上九時至十時到達的概率。

教師應與學生討論累積分佈函數 $\phi(t)$ 的意義。

內容	時間分配	教學建議

離散情況：　$\phi(t)=\sum_{x\leqslant t} f(x)$

連續情況：　$\phi(t)=\int_{-\infty}^{t} f(x)\,dx$

以下的例子，可幫助學生了解上述兩種情況：

1. *離散情況*
   在擲二粒骰子時，其總和大過10的概率為 $1-\phi(10)$。($\phi(10)$ 是其總和等於或少於10的概率。)
2. *連續情況*
   假設 $\phi(a)$ 代表一電燈泡的壽命少於a的概率，於是 $P(X<a)=\phi(a)$，
   $P(a<X<b)=\phi(b)-\phi(a)$ 及 $P(X>a)=1-\phi(a)$

內容	時間分配	教學建議
20.2 預期值及方差	5	教師可與學生重溫平均值及標準差的意義及其具體效用，從而幫助他們了解期望值的概念。教師可用以下的簡單例子，來闡釋此等概念。

例
某人贏取獎金 $x=\$200$ 的概率為 $p=0.01$，則他的機會應值 $px=(\$200)\cdot(0.01)=\$2$。教師可將此概念伸延至X的n個離散值。

教師應定義離散隨機變量及連續隨機變量的期望值。(即 $E(X)=\Sigma px$ 及 $E(X)=\int_{-\infty}^{\infty} x f(x)\,dx$。)

教師亦可討論函數的期望值的定義。以下是兩個定義。

當X是離散隨機變量，　$E[g(X)]=\Sigma p g(x)$

當X是連續隨機變量，　$E[g(X)]=\int_{-\infty}^{\infty} f(x)g(x)\,dx$

在離散隨機變量方面，學生應知道 $E(X)(=\mu)$ 及 $E[(X-\mu)^2](=\mathrm{Var}(X)=\sigma^2)$。教師應特別指出 μ 是集中趨勢的量度，而 σ^2 則是離差的量度。

內容	時間分配	教學建議

以下的例子可供參考。

例

在每一次考試中，甲考生的合格概率是0.4。如果他不合格，他會繼續參加考試，直至合格為止，而每次考試須付費用 \$120。教師可與學生計算此考生考獲合格的預期費用。

與日常生活有關的例子如公平博奕及計算預計盈虧等可幫助學生了解，以下是其中兩個例子。

例一

在一項投資中，甲有0.62的機會賺取 \$5 000，和0.38的機會虧損 \$8 000。

學生可從以下的公式計算 $E(X)=\mu$ 及 $Var(X)=\sigma^2$

$\mu=\$(5\,000)(0.62)+\$(-8\,000)(0.38)=\$60$

$\sigma^2=(5\,000-60)^2(0.62)+\$(-8\,000-60)^2(0.38)$

μ是期望增益。

例二

一部角子機共有四個窗，每一個窗都會顯出紅、橙、黃或藍色，而每種顏色出現的機會是均等的，同時亦不受其他窗所顯示的顏色影響。甲每次投入 \$a，若四個窗所顯示的顏色全不相同，則他會得到獎金 \$5；若四個窗所顯示的顏色完全相同，則獎金為 \$30。在上述兩種情況外他都算輸。\$X 是他在某一次賭博中所得的款項。

在此例中，教師可與學生討論以下問題：

1. 當 $E(X)=0$，這遊戲是一場公平博奕。那麼，遊戲的合理價錢(即 \$a)是多少？
2. 假設 $a=1$，$E(X)$ 及 $Var(X)$ 是多少？

學生須認識在大多數賭博中，$E(X)<0$。教師亦可要求學生計算當甲每次投入的錢加倍時新的 μ 及 σ^2，並找出新與舊參數的關係。若學生程度較佳，教師可着學生估計 $E(Y)$ 的數值其中 \$Y 是甲玩了兩次角子機所贏得的款項。

內容	時間分配	教學建議

在計算一些有關連續隨機變量時，教師應引用例子，清楚列明計算 $E(X)$ 及 $Var(X)$ 的步驟。

例

某一快餐店每天所需的橙汁數量是一個連續隨機變量 X，其概率密度函數 $f(x)$ 的形式如下：

$$f(x)=\begin{cases} ax(b-x) & 0\leqslant x\leqslant 1 \\ 0 & \text{其他} \end{cases}$$

每天所需橙汁的平均數為 0.625 單位。教師可用以下的公式計算 a 及 b：

$$\int_{-\infty}^{\infty} f(x)\,dx=1 \text{ 及 } \int_{-\infty}^{\infty} x\,f(x)\,dx=0.625$$

盛載橙汁的容器總容量為 0.8 單位，每天早上，此容器均給注滿。

假設在某一天，橙汁存量供不應求的概率為 p，p 可根據以下公式求得：

$$p=1-\phi(0.8)=\int_{0.8}^{1} f(x)\,dx$$

對一些程度較佳的學生而言，教師應引導他們去證明以下兩條公式：

$E[a\,g(x)+b]=a\,E[g(x)]+b$ 及 $Var(aX+b)=a^2\,Var(X)$ 其中 a 與 b 均是常數。

此外，一般學生應能證明：

$E[g(x)+h(x)]=E[g(x)]+E[h(x)]$ 及 $Var(X)=E(X^2)-[E(X)]^2=E(X^2)-\mu^2$

教師可用以下的例子來解釋上述公式的應用。

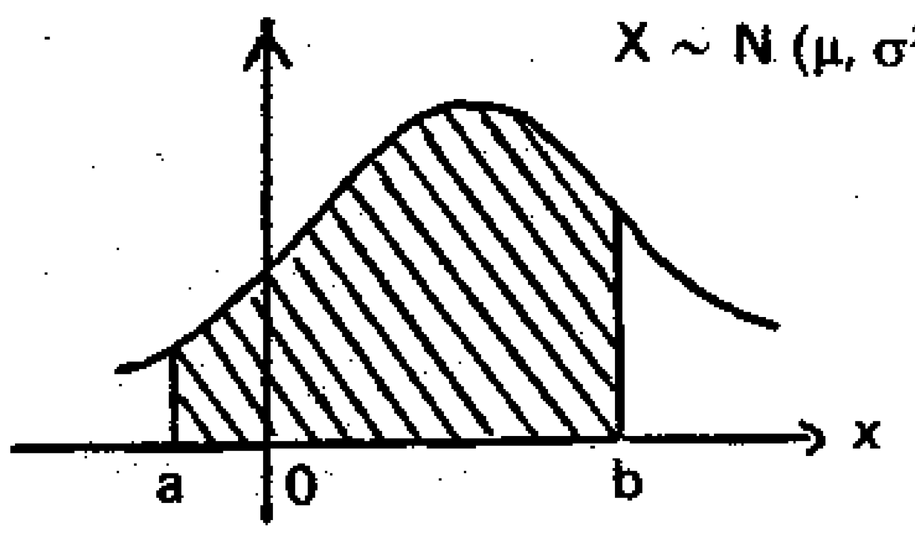
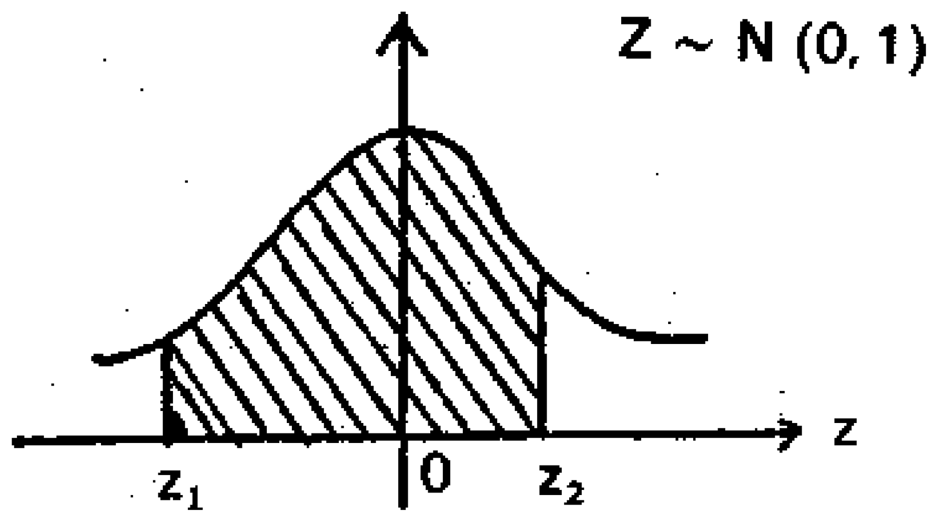
例

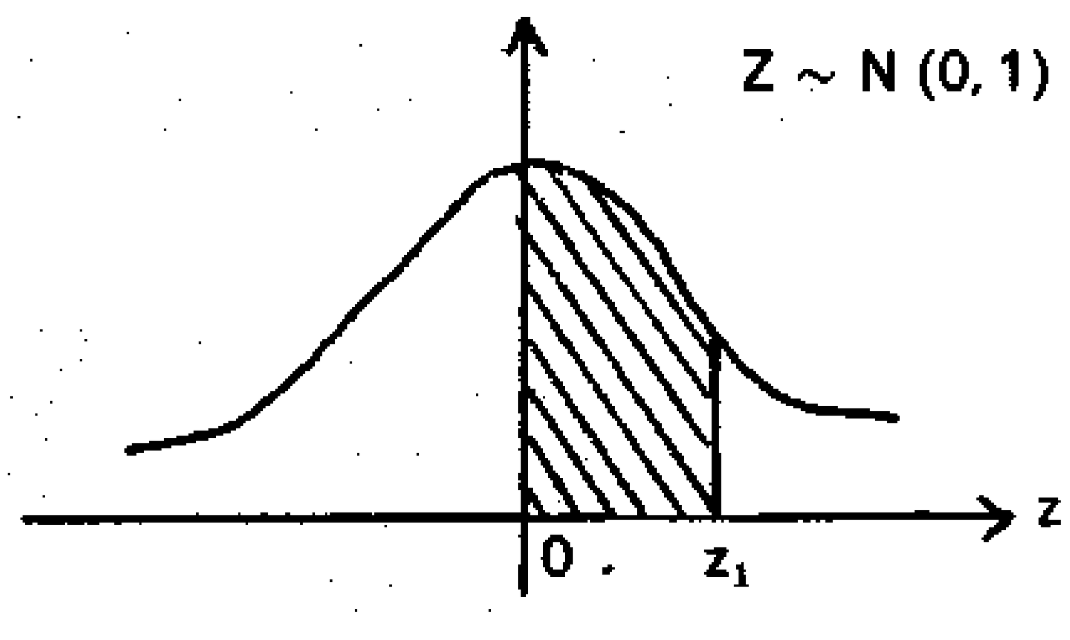
假設 $Z=2X^2-3X+5$，其中 X 是一隨機變量，其平均值為 μ，方差為 σ^2。$E(Z)$ 可從以下的公式計算出來。

$$\begin{aligned} E(Z)&=E(2X^2-3X+5)=E(2X^2)-E(3X)+E(5) \\ &=2(\mu^2+\sigma^2)-3\mu+5 \end{aligned}$$

內容	時間分配	教學建議
20.3 二項分佈 (*a*) 伯努利試驗 二項概率	7	教師可引用已熟習的擲毫例子來介紹伯努利試驗，並強調在伯努利試驗中，只有兩種可能的結果。 重覆的伯努利試驗，在概率學及統計學上担當一個重要的角色，尤其是當此兩種可能結果，在每一次試驗中的概率是相等的時候。學生應知道在n次試驗後，共有r次成功的概率為： $P(r\text{次成功})=C^n_r p^r q^{n-r}$ 教師可引用更多的日常例子，來闡釋二項概率。 *例一* 擲一粒骰子n次，為使得到至少一個「六」的機會大過0.99，n須符合以下的不等式： $1-(\frac{5}{6})^n>0.99$ *例二* 隨機將r個球分佈在n格內。教師可要求學生計算在某一指定格內放存k個球($k\leqslant r$)的概率P_k。 在此例中，學生須知道$P_k=P$(在r次伯努利試驗中有k次成功事件)，而$p=\frac{1}{n}$。
(*b*) 二項分佈		此時，教師可介紹二項分佈可視作一個重覆的伯努利試驗，而其成功的概率是相等的。此外亦可介紹這分佈的記號$B(n, p)$。 學生須認識二項分佈的公式，$E(X)=np$及$Var(X)=npq$。 教師可利用不同n及p的數值，來顯示二項分佈的概率圖表。學生應可觀察到當$p=q=0.5$時，此圖表是對稱的。對能力較高的學生，教師亦可討論二項分佈的眾數。
(*c*) 應用		二項概率分佈可用於很多實際的事件中，以下是三個例子。 *例一* 在某項測驗中，共有4條選擇題，每題均有5項選擇。甲參加了是項測驗；以他所學的知識，他每條選中正確答案的概率為0.7；若他只是瞎猜，則選中正確答案的概率是0.2。甲做了全部問題。

內容	時間分配	教學建議
		在上述例子中，甲只知道其中3題的正確答案的概率是$C^4_3(0.7)^3(0.3)$。 由於甲只是瞎猜也可能得到正確答案，所以P(正確答案)=p可在以下兩種情況計算出： (*a*) 他確實懂得此題； (*b*) 他只是瞎猜得到的答案。 假設X是所選正確答案的數目，教師可要求學生計算$E(X)$ $(=4p)$, $Var(X)$ $(=4p(1-p))$及$P(X=1)$ $(=C^4_1(p)^1(1-p)^3)$。 教師可提出以下問題： 1. 答對一題得兩分，答錯一題扣一分。假設甲得Y分，計算$E(Y)$及$Var(Y)$。 2. 假設甲只知道其中3題的正確答案，他得滿分的概率為何？ 3. 假設他只選中一個正確答案，他由於瞎猜而得出此答案的概率為何？ *例二* 一批燈泡中有5%是次貨。現要根據以下的規則，測驗一批燈泡。 (*a*) 測試其中10個燈泡的樣本。 (i) 若有兩個或以上的燈泡是次貨，則退回整批貨。 (ii) 若在此樣本中沒有次貨，則接受整批貨。 (iii) 若只有一個次貨，則進行以下(*b*)的測試。 (*b*) 再測試另外10個燈泡的樣本。若在今次測試中沒有次貨，則接受整批貨，否則退回。 若有X個的燈泡接受測試，則要找出$P(X=20)=10(0.95)^9(0.05)$及$P(X=10)=1-10(0.95)^9(0.05)$並不困難。教師亦可要求學生計算$E(X)$及$Var(X)$。 *例三* 一機器所生產的貨物有10%是次貨。所有貨物先放在一大盒內，由每盒抽出n件貨物測試。若沒有次貨，則接受整批貨，否則退回。 為使退貨的概率最小是0.95，n之最低值必須符合$(0.95)^n<1-0.95$。若$n=10$，則P(接受貨物)$=(0.9)^{10}=p$。 在8盒被測試的貨物中，有5盒被退回的概率是$C^8_3 p^3(1-p)^5$。

內容	時間分配	教學建議
20.4 常態分佈 (*a*) 基本定義	10	常態分佈是一個十分重要的連續概率分佈例子。教師只須寫下 p.d.f., f(x) 的公式，但不用詳加解釋。 $$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$ 學生只須認識 $E(X)=\mu$ 及 $Var(X)=\sigma^2$，而無須予以證明。 教師亦可與學生討論常態分佈被廣泛使用的原因。 1. 容易運用； 2. 可作為其他分佈的近似分佈。 教師可介紹由不同 μ 和 σ 所作的圖表，學生應認識所有圖表均是鐘形，並對 $x=\mu$ 對稱。教師可介紹 $N(\mu，\sigma^2)$ 是常態分佈的記號，其中 μ 為平均值及 σ^2 為方差。
(*b*) 標準常態曲綫和常態表的使用		常態分佈運算跟 μ 及 σ 有莫大關係。學生應發現如以不同的參數來表列不同的常態分佈概率函數並不容易。因此，有必要以標準單位 $Z=\frac{X-\mu}{\sigma}$ 作變換。學生應不難得知 $E(Z)=0$, $Var(Z)=1$ 及 $$P(a<X<b)=P(\frac{a-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{b-\mu}{\sigma})$$ $$=P(z_1<Z<z_2)$$ 下列圖表，可幫助說明以上的概念。  有陰影部分的面積相等

內容	時間分配	教學建議
		在下圖中，陰影部分的面積是 $P(0<Z<z_1)$  以不同的 z_1 值計算出來的面積以表形式列出，稱為常態分佈表(此表只列到 $z_1=3.59$ 的數值)。學生須多加練習，才可正確地使用此圖表。 *例* X是 N(8, 4) $$P(6<X<9)=P\left(\frac{6-8}{2}<\frac{X-8}{2}<\frac{9-8}{2}\right)$$ $$=P(-1<Z<0.5)$$ $$=P(0<Z<1)+P(0<Z<0.5)$$ $P(X>9)=P(Z>0.5)=0.5-P(0<Z<0.5)$ 若 $P(X<k)=0.87$，則可用綫性插值法準確地計算 k 值。 教師可提醒學生在解決此類問題時，可利用對稱性及互補概率。此外，若 z_1 值超過三位有效數字時，應採用綫性插值法。
(*c*) 應用		在日常應用中，標準常態分佈是非常重要的。教師應提供足夠的例子說明。下列的例子可供參考。

內容	時間分配	教學建議
		例一 某生產商用機器生產電阻。他發現其中10%的電阻少於95Ω，同時，其中20%是超過110Ω。電阻X是屬於常態分佈。 μ及σ可由以下兩條公式求得： $P(X<95)=P\left(Z<\frac{95-\mu}{\sigma}\right)=0.1$ $P(X>110)=P\left(Z>\frac{110-\mu}{\sigma}\right)=0.2$ *例二* X是一支杆的長度，它是一個常態分佈的隨機變量，其平均值為μ，方差為1。若X不符合某些標準，生產商便會蒙受損失。假設利潤M(每支杆計)是X的函數如下： $M=\begin{cases} 3 & 8\leqslant x\leqslant 10 \\ -1 & x<8 \\ -5 & x>10 \end{cases}$ 預期的利潤E(M)為$E(M)=8\phi(10-\mu)-4\phi(8-\mu)-5$， 其中$\phi(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$是累積概率函數。 假設可調節生產過程而使$\mu$的值改變，則相對於最大利潤的$\mu$值可從E(M)對$\mu$的導數中求得。 *例三* 一工廠生產樽裝汽水，每樽容量為1.25L。但是，由於自動入樽機的偏差，每樽所含的汽水容量會隨常態分佈而改變。據觀察所得，其中15%的汽水容量少於1.25L，而其中10%容量超過1.30L。 學生應可求得容量分佈的平均值μ和標準偏差σ。 生產一樽有x L汽水的成本(以仙計)為$C=36+62x+5x^2$其中x是以上分佈的隨機變量。 每樽汽水的預期成本是E(C)，其中$E(C)=36+62\mu+5(\mu^2+\sigma^2)$。 20 000樽汽水的預期成本是20 000 E(C)。

內容	時間分配	教學建議
(*d*) 逼近常態分佈的二項分佈		學生應清楚知道只有當n的數值大時，才可用常態近似法來計算二項概率。在此情況下，平均值及方差分別為np及npq。此外，教師亦須提醒學生在這近似計算時必須作末端的校正。 *例* 擲一枚硬幣400次。 若X是所得正面的次數，則X是B(400, 0.5)。用常態近似法，它是N(200, 100)。 $P(195\leqslant X\leqslant 210)=P\left(\frac{194.5-200}{10}<Z<\frac{210.5-200}{10}\right)$ 學生會有興趣知道： $P(195\leqslant X\leqslant 210)\neq P(195<X<210)$
20.5 獨立常態變量的綫性組合	6	學生應認識到獨立常態變量的純量倍數和亦是常態。由此，學生不難推論： 若X及Y是兩個獨立常態變量，且$X\sim N(\mu_1, \sigma_1^2)$及$Y\sim N(\mu_2, \sigma_2^2)$，則$aX+bY\sim N(a\mu_1+b\mu_2, a^2\sigma_1^2+b^2\sigma_2^2)$而a、b是任意實數。 以上結果可輕易伸展至n個獨立常態變量。教師可舉例說明以上事實在日常生活中的應用。 *例一* 蛋糕以六個一袋出售。每一塊蛋糕的質量是常態變量而其平均值為25 g及標準差為2 g。包裝蛋糕的物料質量亦是常態變量而平均值是30 g及標準差是4 g。求每袋蛋糕總質量的分佈及由此求出每袋蛋糕的總質量小於142 g的概率。 *例二* 平裝書的厚度，A cm，是常態分佈，其平均值為2cm及方差為0.63cm²。精裝書的厚度，B cm，亦是常態分佈而其平均值為5 cm及方差為1.42 cm²。學生不難求得$X=2A-B$的分佈；由此，可用常態分佈表求得一隨機選擇的平裝書厚度小於一隨機選擇的精裝書厚度的一半的概率。
	34	