Mathematical Statistics

Chapter 2: Random Variables

Huei-Wen Teng

Dept Information Management and Finance
National Yang Ming Chiao Tung University

Type of data

- Qualitative: multiple categorical
- Quantitative: discrete, continuous

Identical expressions for a random variable $X \sim F(\theta)$

- Notation: $X \sim \mathrm{Name}(\theta)$ with θ the parameter
- Two types:
  - discrete: probability mass function (pmf), $p(x)$
  - continuous: probability density function (pdf), $f(x)$
- Cumulative distribution function (cdf): $F(x)$
- Moment generating function (mgf): $M_X(t)$.
- Quantitative: discrete, continuous

$$F(x) = P\left(X \leq x\right)$$

2.1 Discrete Random Variables

Original Definition of RV

- Random variable (RV) is a function from Ω to the real numbers.

EX: Toss a fair coin with outcome head (H) and tail (T).

- The sample space is $\Omega = \{H, T\}$.
- Set a RV: $\begin{cases} X(H) = 1, \\ X(T) = 0. \end{cases}$
- Then, the sample space of X is $\Omega = \{0,1\}$.
- Then, we have $\begin{cases} P(X = 1) = P(H) = 1/2, \\ P(X = 0) = P(T) = 1/2. \end{cases}$

Definition of a discrete RV

- A discrete random variable is a random variable that can take on only a finite or at most a countably infinite number of values.
  - The probability mass function or the frequency function is
    $p(x) = P\left(X = x\right)$
  - The cumulative distribution function (cdf) of a random variable is defined as
    $F(x) = P\left(X \leq x\right), -\infty < x < \infty$

Remark

- Once you know a pmf of a random variable, you can derive its cdf by definition:

$$F(x) = \sum_{k=-\infty}^{x} P\left(X = k\right)$$

- Once you know a cdf of a random variable, you can find its pmf by:

$$P\left(X = k\right) = F(k) - F(k-1)\,.$$

Bernoulli random variables, $X \sim \mathrm{Bern}(p)$

- A Bernoulli random variable takes on only two values: 0 and 1, with probabilities $1 - p$ and p, respectively. Its frequency function is

$$\begin{cases} p(1) & = & p, \\ p(0) & = & 1 - p, \\ p(x) & = & 0, \text{ if } x \neq 0 \text{ and } x \neq 1. \end{cases}$$

Bernoulli random variables, $X \sim \mathrm{Bern}(p)$

An alternative and sometimes useful representation is

$$p(x) = \begin{cases} p^x(1-p)^{1-x}, & \text{if } x = 0, \text{ or } x = 1, \\ 0, & \text{otherwise}\,. \end{cases}$$

The Binomial distribution, $X \sim \text{Binomial}(n, p)$

- Suppose that n independent experiments, or trials, are performed, where n is a fixed number, and that each experiment results in a "success" with probability p and a failure with probability $(1-p)$. The total number of successes, X, is a binomial random variable with parameters n and p.

$$p(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, 1, 2, \cdots, n\,. \\ 0, \text{ otherwise} \end{cases}$$

The Geometric distribution, $X \sim \mathrm{Geo}(p)$

- The Geometric distribution is constructed from independent Bernoulli trials but from an infinite sequence. On each trial, a success occurs with probability p, and X is the total number of trials up to and including the first success. So,

$$p(k) = P\left(X = k\right) = (1-p)^{k-1}p,\ k = 1, 2, 3,\ \cdots$$

The negative binomial distribution, $X \sim \text{negBinomial}(r, p)$

- The negative binomial distribution arises as a generalization of the geometric distribution. Suppose that a sequence of independent trials, each with probability of success p, is performed until there are r successes in all; let X denote the total number of trials.

$$p(k) = P\left(X = k\right) = \binom{k-1}{r-1} p^r (1-p)^{k-r},\ k = r, r+1, r+2, \cdots$$

I will skip the Hypergeometric distribution, because it is difficult to memorize its formula. Interestingly, it is easier to find the pmf directly.

The Hypergeometric distribution, $X \sim \mathrm{HyperG}(r, n, m)$

- Suppose that an urn contains n balls, of which r are black, and $n - r$ are white. Let X denote the number of black balls drawn when taking m balls without replacement. The parameters need to satisfy $m \leq n$, and $r \leq n$.

And,

$$P(X = k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}},$$

for $k = \max(m - (n - r), 0), 1, 2, \cdots, \min(m, r)$.

The support of $\mathrm{HyperG}(r, n, m)$: $x_{\min}$

The possible minimum value of X, denoted as $x_{\min}$

- If there are more than or exactly m white balls, i.e., $(n - r) \geq m$, $x_{\min} = 0$.
- If there are less than m white balls, i.e., $n - r < m$, $x_{\min} = m - (n - r)$.
- Rewrite these two conditions with one formula, $x_{\min} = \max\big(m - (n - r), 0\big)$.

The support of $\mathrm{HyperG}(r, n, m)$: $x_{\max}$

The possible minimum value of , denoted as

- If there are more than or exactly m white balls, i.e.,

  $r \geq m$, $x_{\max} = m$.

- If there are less than m white balls, i.e., $r < m$, $x_{\max} = r$.
- Rewrite these two conditions with one formula,

  $x_{\max} = \min(m, r)$.

Before the Poisson Distribution

- Recall the Taylor expansion:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

- Use different dummy variable:

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

- Divide the above formula by e^λ,

The Poisson Distribution

The Poisson frequency function with parameter is

$$P(X=k)=e^{-\lambda}\frac{\lambda^{k}}{k!},\ k=0,1,2,\cdots$$

The Poisson Distribution (Skipped)

It can be derived as the limit of a binomial distribution when $n \to \infty$ and $np = \lambda$.

Set $np = \lambda$,

$$\begin{aligned} p(k) &= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \frac{n!}{(n-k)!} \frac{1}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &\overset{n \to \infty}{=} \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

2.2 Continuous Random Variables

Continuous random variable

- A continuous random variable takes on a continuum of values rather than a finite or countably infinite number.
- For a continuous random variable, the role of the frequency function is taken by a density function have the properties that
  - $f(x) \geq 0$
  - f is piecewise continuous
  - $\int_{-\infty}^{\infty} f(x)\, dx = 1.$

Continuous random variable

If X is a random variable with a density function f, then for any $a < b$, the probability that X falls in the interval (a, b) is the area under the density function between a and b:

$$P\big(a < X < b\big) = \int_a^b f(x)\,dx\,.$$

f is also known as probability density function (pdf).

Continuous random variable. Continued.

The cumulative distribution F hence links to the density function by

$$F(x) = P\big(X \leq x\big) = \int_{-\infty}^{x} f(u)\,du\,.$$

Remark

- Once you know a pdf of a random variable, you can derive its cdf by definition:

$$F(x) = \int_{-\infty}^{x} f(u)\, du\,.$$

- Once you know a cdf of a random variable, you can find its pdf by differentiation:

$$f(x) = \frac{d}{dx}F(x) = \frac{d}{dx}\int_{-\infty}^{x} f(u)\, du\,.$$

The Uniform random variable, $X \sim U(a, b)$

The uniform random variable on the interval $[a, b]$ has a density function

$$f(x) = \begin{cases} 1/(b-a), & a \leq x \leq b, \\ 0, & x < a \text{ or } x > b\,. \end{cases}$$

Review on calculus. Fundamental theorem of calculus. First part.

Let f be a continuous real-valued function defined on a closed interval $[a, b]$.

Let F be the function defined, for all x in $[a, b]$, by

$$F(x) = \int_a^x f(t)\, dt\,.$$

Then, F is uniformly continuous on $[a, b]$, differentiable on the open interval (a, b), and

$$F'(x) = f(x)$$

for all x in (a, b). Source: https://en.wikipedia.org/wiki/Fundamental_theorem_of_calculus

Example

- $X \sim U(0,1)$. Find $P(-0.2 < X < 0.3)$.
- Solution.

$$
\begin{aligned}
P(-0.2 < X < 0.3) &= \int_{-0.2}^{0.3} f(x)dx \\
&= \int_{0}^{0.3} 1dx \\
&= 0.3 \times 1 = 0.3
\end{aligned}
$$

Review on calculus. Fundamental theorem of calculus. Second part.

Let f be a real-valued function on a closed interval $[a, b]$ and F an antiderivative of f in $[a, b]$:

$$F'(x) = f(x) .$$

If f is Riemann integrable on $[a, b]$ then

$$\int_a^b f(x)\, dx = F(b) - F(a) .$$

Must know in calculus

- Integration
  - Drawing figures
  - Integration for integrand: polynomial and exponential
  - Integration by part
  - Change of variable
- Differentiation
  - Differentiation for polynomial, exponential
  - Chain rule
  - Leibniz rule (optional!)

The Exponential random variable, $X \sim \mathrm{Exp}(\lambda)$

- Here, λ is the rate parameter.
- The exponential density function is

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

The Exponential random variable. Continued.

The cumulative distribution is easily found:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

The memory-less property

- Suppose $T \sim \mathrm{Exp}(\lambda)$.
- The exponential distribution has the memory-less property:

$$
\begin{aligned}
P\Big(T > t + s \,\Big|\, T > s\Big) &= \frac{P\big(T > t + s \text{ and } T > s\big)}{P(T > s)} \\
&= \frac{P\big(T > t + s\big)}{P(T > s)} \\
&= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} \\
&= e^{-\lambda t} \\
&= P(T > t)\,.
\end{aligned}
$$

The Gamma random variable, $X \sim G(\alpha, \lambda)$

- The gamma density function is

$$g(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \; t \geq 0.$$

where α is called a shape parameter and λ is called the rate parameter.

- The gamma function, $\Gamma(t)$, is defined as

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} \, du, \; x > 0.$$

- Special cases

$$G(1, \lambda) \sim \text{Exp}(\lambda) \,.$$

Gamma densities

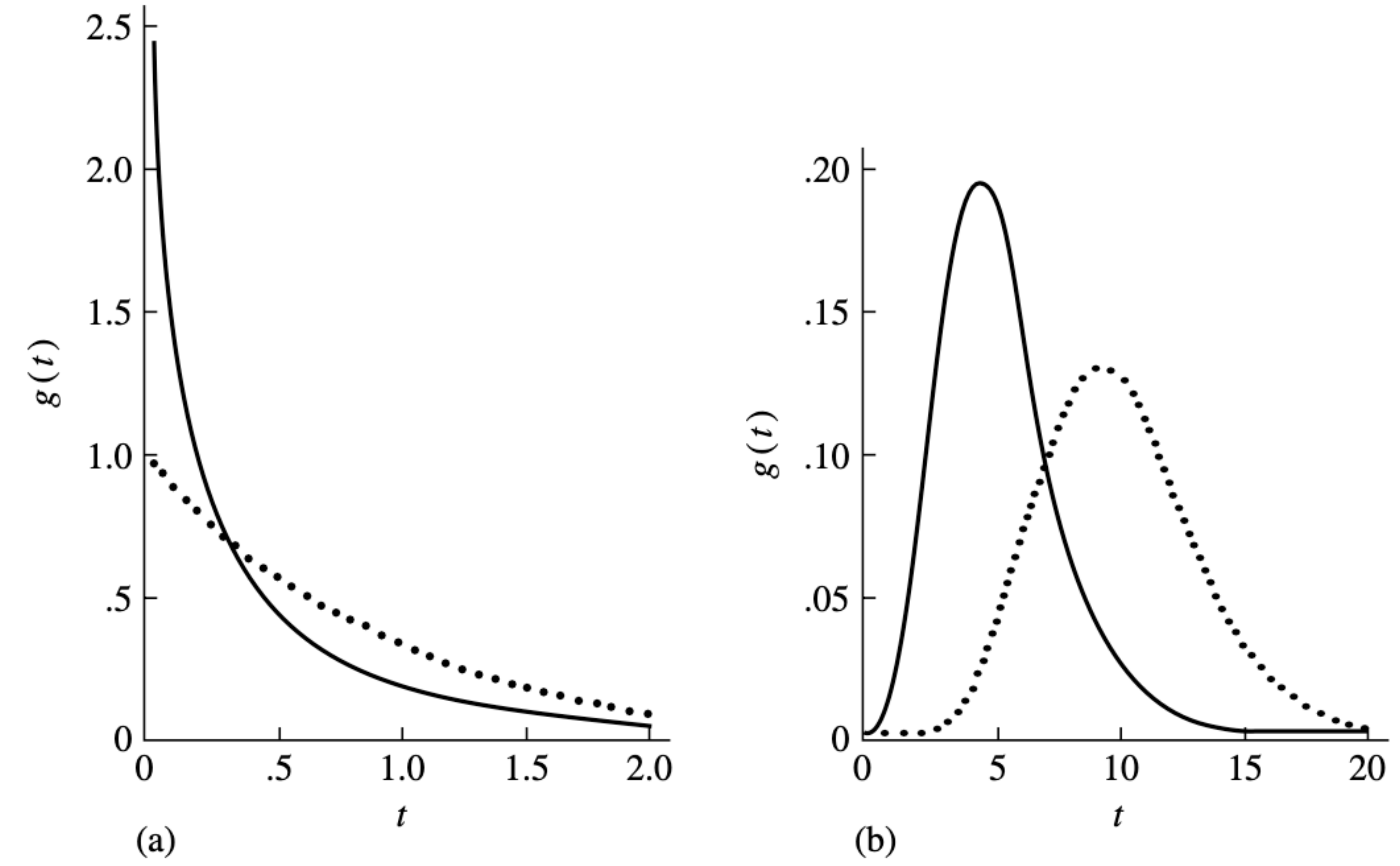


FIGURE **2.11** Gamma densities, (a) $\alpha = .5$ (solid) and $\alpha = 1$ (dotted) and (b) $\alpha = 5$ (solid) and $\alpha = 10$ (dotted); $\lambda = 1$ in all cases.

Gamma densities

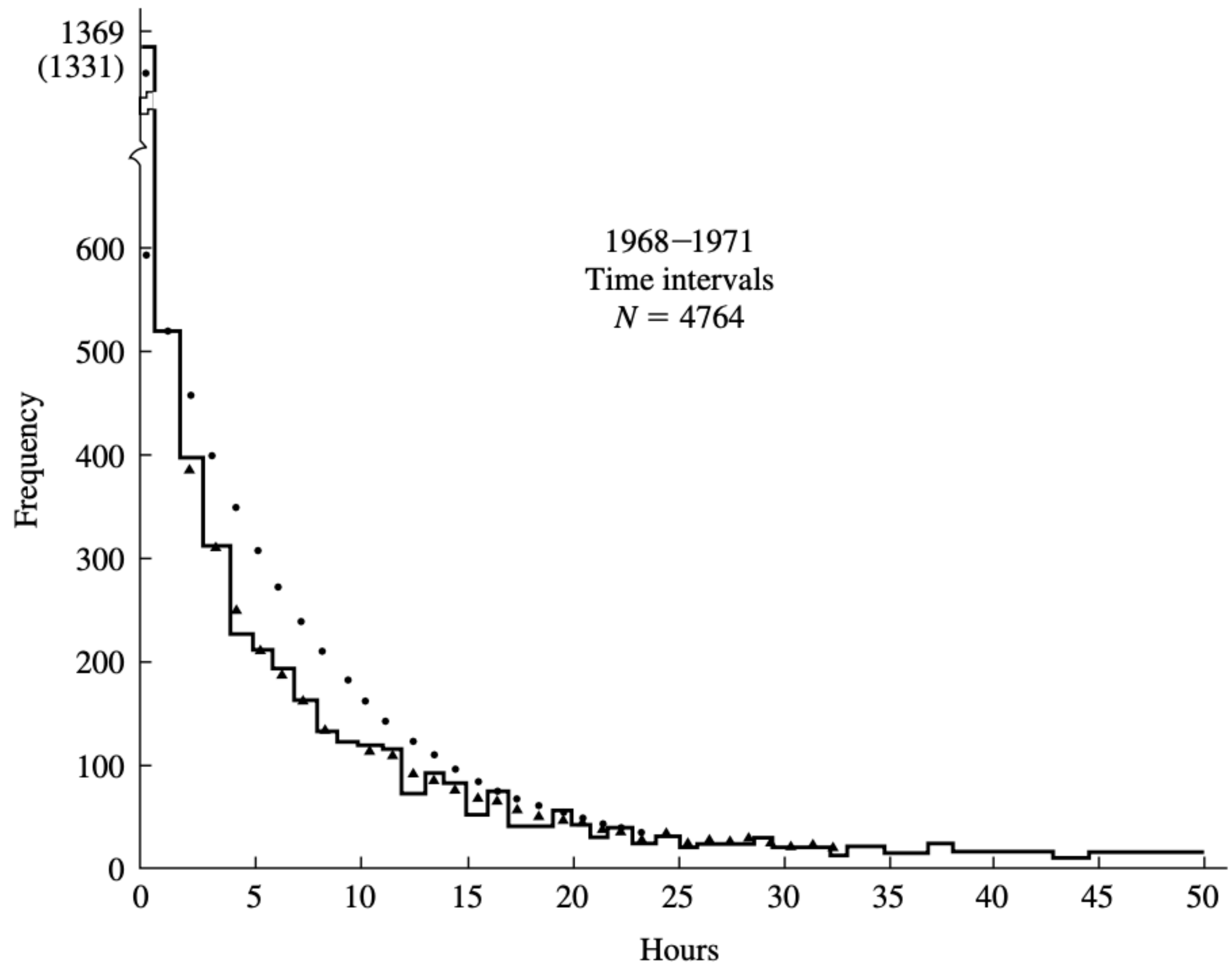


FIGURE **2.12** Fit of gamma density (triangles) and of exponential density (circles) to times between microearthquakes.

Notes on the Gamma function

The gamma function $\Gamma(z) = \int_0^\infty x^{z-1} \exp(-x)\, dx.$

1. $\Gamma(1/2) = \sqrt{\pi}$
2. $\Gamma(1) = 1$
3. $\Gamma(z+1) = z\Gamma(z)$
4. $\Gamma(n) = (n-1)!$

Proof: Homework!

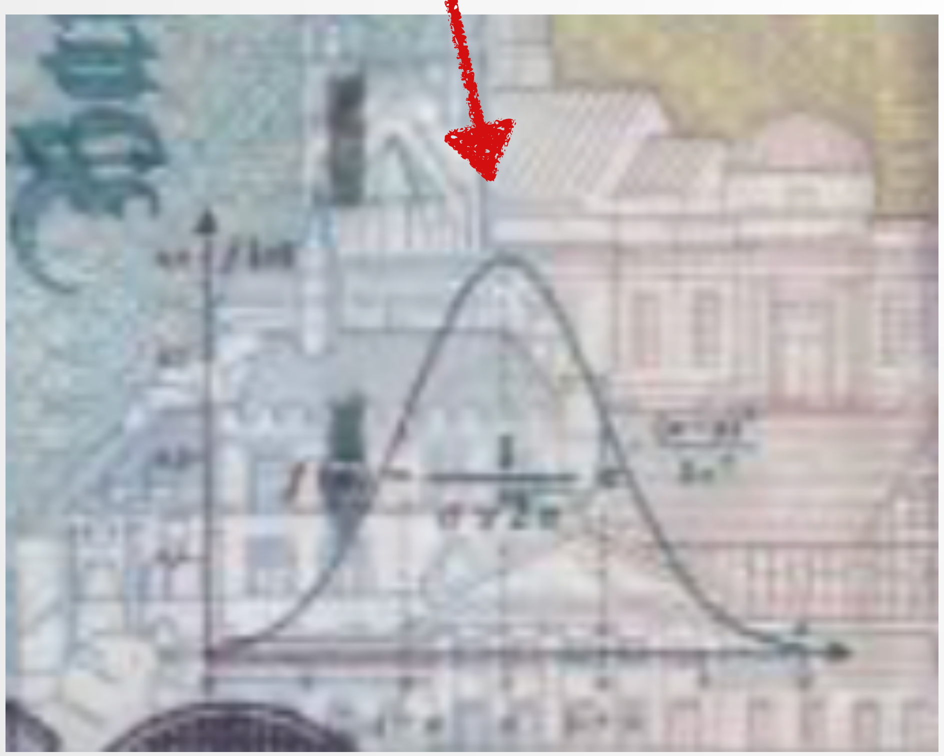
The Normal random variable, $X \sim N(\mu, \sigma^2)$

- The parameters μ and σ are called the mean and standard deviation.
- The density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-(x-\mu)^2/2\sigma^2 \Big), \quad -\infty < x < \infty \,.$$

- When $\mu = 0$ and $\sigma = 1$, it is called the standard normal random variable, usually denoted as Z.
  - The density is denoted by $\phi(\,\cdot\,)$ and
  - Its cdf is denoted by $\Phi(\,\cdot\,)$.

Normal densities



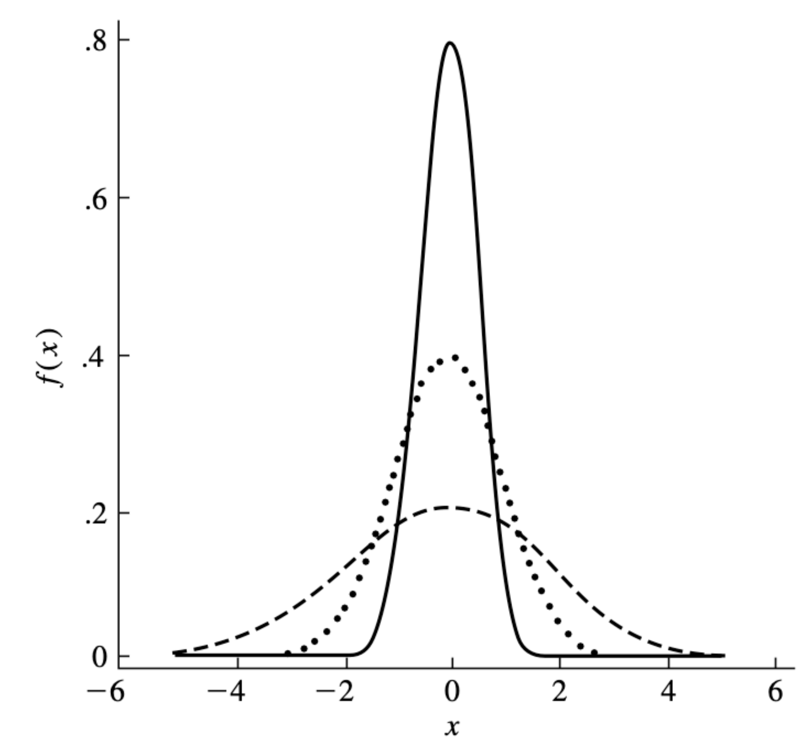


FIGURE **2.13** Normal densities, $\mu = 0$ and $\sigma = .5$ (solid), $\mu = 0$ and $\sigma = 1$ (dotted), and $\mu = 0$ and $\sigma = 2$ (dashed).

Example

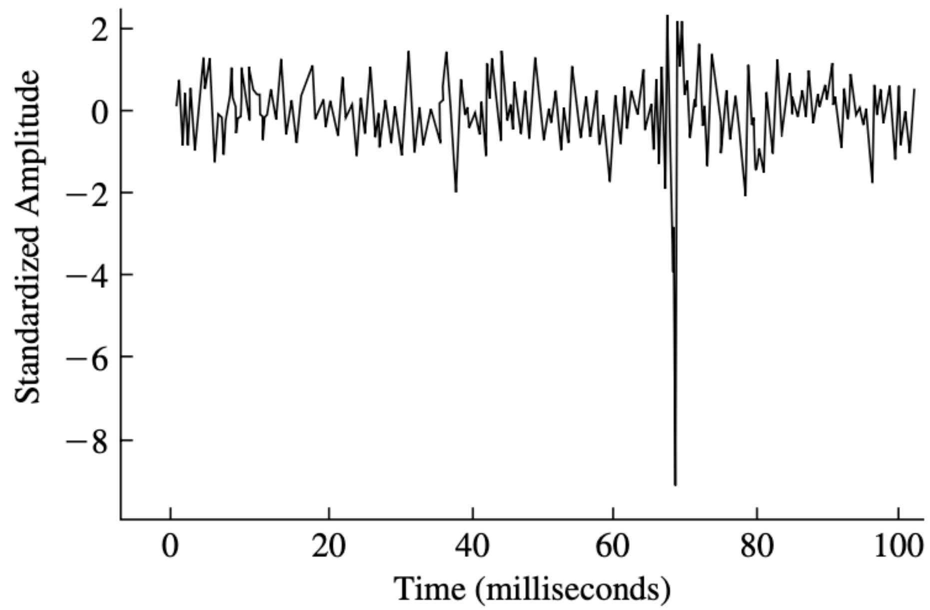


FIGURE 2.14 A record of undersea noise containing a large burst.

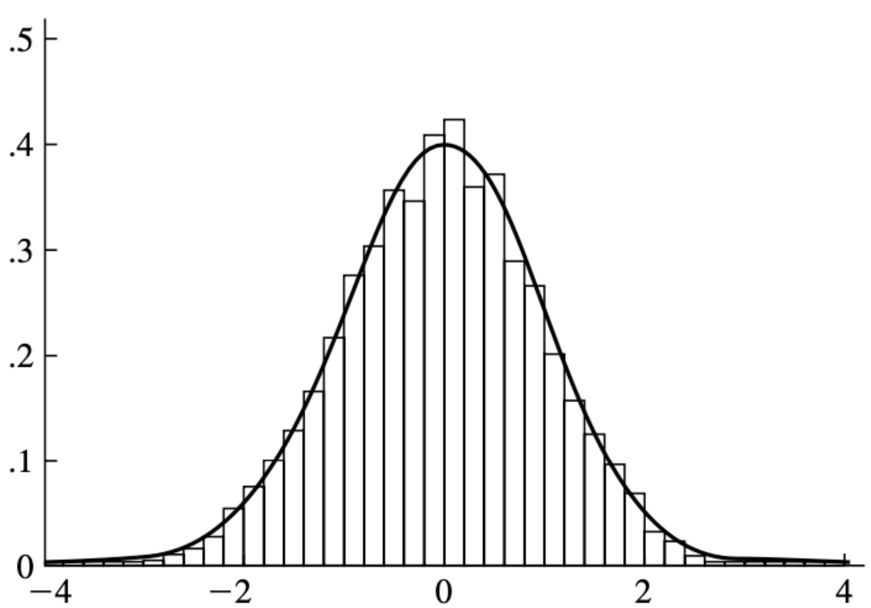


FIGURE 2.15 A histogram from a "quiet" period of undersea noise with a fitted normal density.

Veitch and Wilks (1985) studied recordings of Arctic undersea noise and characterized the noise as a mixture of a Gaussian component and occa- sional large-amplitude bursts.
Figure 2.14 is a trace of one recording that includes a burst.
Figure 2.15 shows a Gaussian distribution fit to observations from a "quiet" (nonbursty) period of this noise.

Example

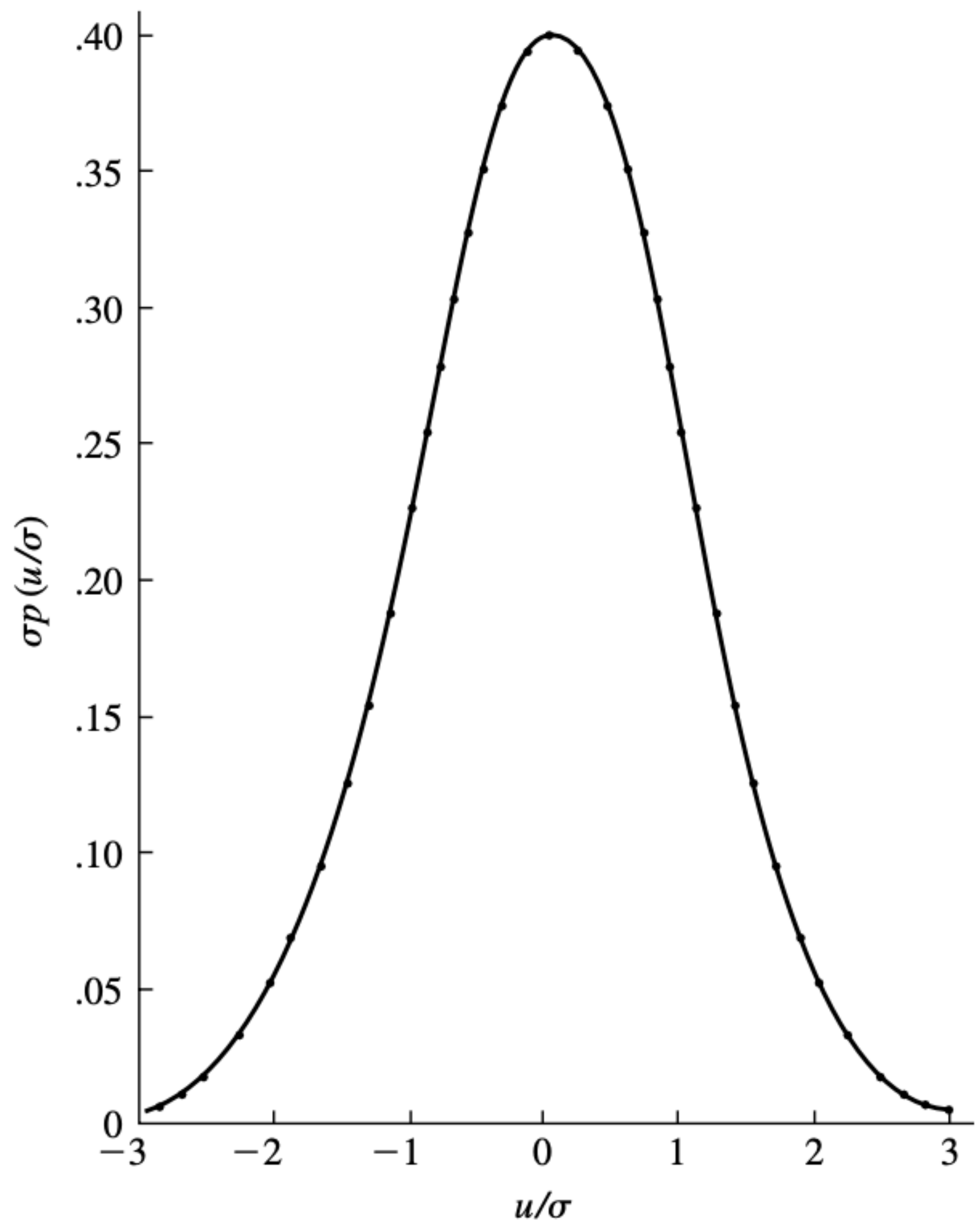


FIGURE **2.16** A normal density (solid line) fit to 409,600 measurements of one component of the velocity of a turbulent wind flow. The dots show the values from a histogram.

Example

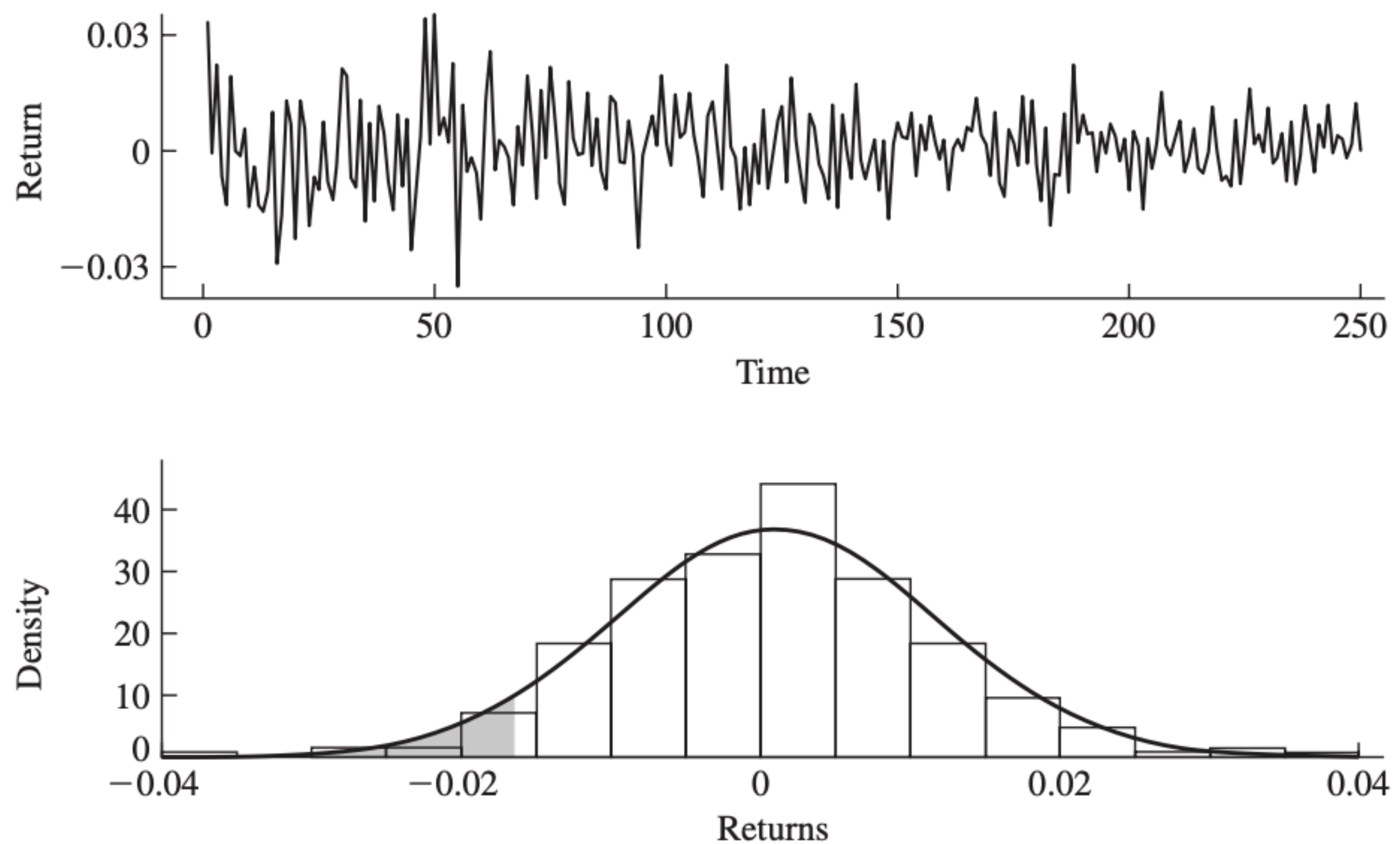


FIGURE **2.17** Returns on the S&P 500 during 2003 (top panel) and a normal curve fitted to their histogram (bottom panel). The region area to the left of the 0.05 quantile is shaded.

The Beta Density

- $$f(u) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{(a-1)}(1-u)^{(b-1)}, \; 0 \leq u \leq 1.$$

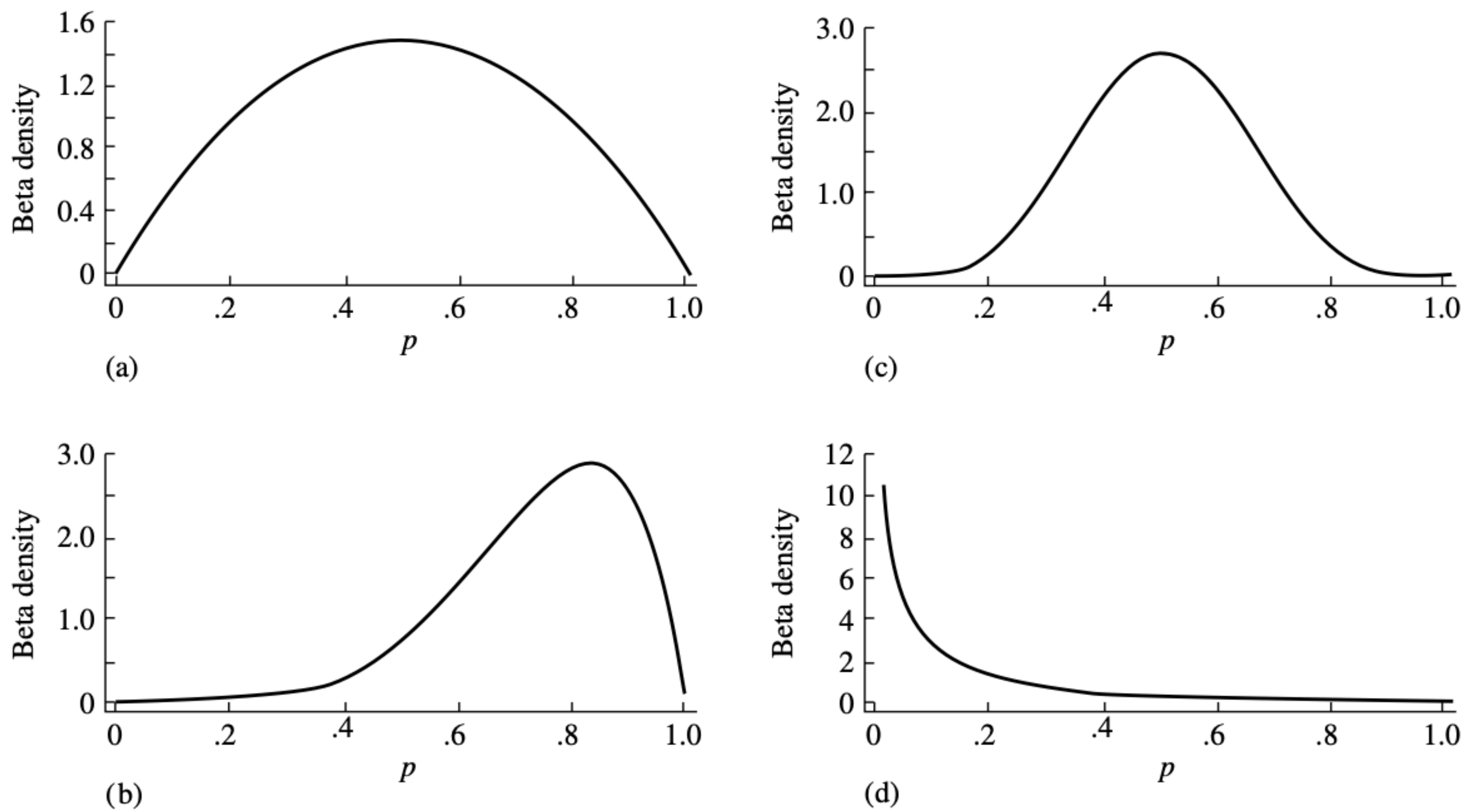


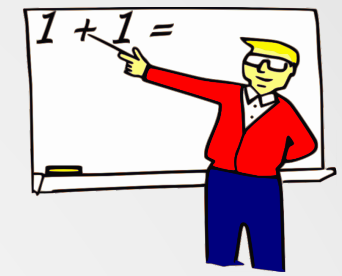
FIGURE **2.18** Beta density functions for various values of a and b: (a) $a = 2, b = 2$; (b) $a = 6, b = 2$; (c) $a = 6, b = 6$; and (d) $a = .5, b = 4$.

2.3 Functions of a Random Variable

A simple example (in-class exercise)!

- X is a discrete random variable with pmf:
- $P(X=-1)=0.1$, $P(X=0)=0.2$, $P(X=1)=0.3$, $P(X=2)=0.4$.
- Find the pmf of Y if $Y=X^2$.

A simple example! ii



- Solution

$X = x$	$p(x)$	$Y = X^2$
-1	0.1	1
0	0.2	0
1	0.3	1
2	0.4	4

Therefore, the pmf of Y is:

$Y = y$	$p(y)$
0	0.2
1	0.4
4	0.4

A S.O.P. to find the pdf of RVs

- Suppose that a random variable X has a density function $f(x)$.
- Need to find the density of $Y = g(X)$ for some given function
- The steps are
  - Draw (a) the pdf $f(x)$ and (b) $y = g(x)$ to find out possible values of Y.
  - Find $F_Y(y)$;
  - Obtain $f_Y(y) = \dfrac{dF_Y(y)}{d(y)}$.
  - Recognize the distribution of a RV if there are any.

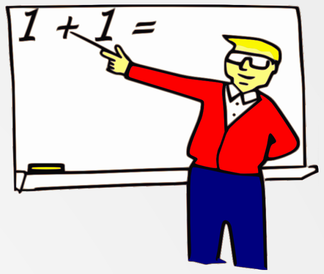
A simple example (in-class exercise)!

- Find the distribution of $Y = 3 + 2X$, where $X \sim U(0,1)$

Proposition A. i

- Assume $X \sim N(\mu, \sigma^2)$.
- Let $Y = aX + b$, then $Y \sim N\left(a\mu + b, a^2\sigma^2\right)$.

Proof of Proposition A.



1. Possible outcome of Y is $\mathbb{R}$.
2. Suppose $a>0$.

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(aX + b \le y) \\
&= P\big(X \le (y-b)/a\big) \\
&= F_X\big((y-b)/a\big)
\end{aligned}
$$

Proof of Proposition A.

3. To get the pdf, we differentiate
4. We see that.

$$Y \sim N\left(b + a\mu, a^2\sigma^2\right)$$

5. (When $a < 0$, the proof is similar.)

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\frac{y-b}{a}\right)$$

$$= \left[\frac{dF_X\left(\frac{y-b}{a}\right)}{d\left(\frac{y-b}{a}\right)}\right]\frac{d\left(\frac{y-b}{a}\right)}{dy}$$

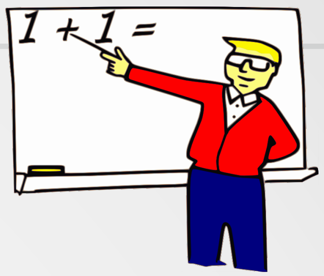
$$= f_X\left(\frac{y-b}{a}\right)\frac{1}{a}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}\right)\cdot\frac{1}{a}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2 a^2}}\exp\left(-\frac{\left(y-a\mu-b\right)^2}{2a^2\sigma^2}\right)$$

Example i

- Find the density of $X = Z^2$, where $Z \sim N(0,1)$.



Solution

- Draw $x = z^2$ and the pdf of Z, to find the possible values of

$$X : \{x : x \geq 0\}$$

- Get $F_X(x)$. For $x \geq 0$,

$$\begin{aligned} F_X(x) &= P(X \leq x) = P\left(Z^2 \leq x\right) \\ &= P\left(-\sqrt{x} \leq Z \leq \sqrt{x}\right) \\ &= \Phi\left(\sqrt{x}\right) - \Phi\left(-\sqrt{x}\right) \end{aligned}$$

Solution

- Get $f_X(x)$. For $x \geq 0$,

$$f_X(x) = \phi\left(\sqrt{x}\right)\frac{1}{2}x^{-\frac{1}{2}} - \phi\left(-\sqrt{x}\right)\frac{-1}{2}x^{-\frac{1}{2}}$$

$$= x^{-\frac{1}{2}}\phi(\sqrt{x})$$

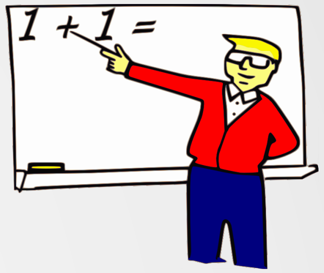
$$= x^{-\frac{1}{2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x}{2}\right)$$

$$= \frac{\frac{1}{2}^{\frac{1}{2}}}{\Gamma(1/2)}x^{\frac{1}{2}-1}\exp\left(-\frac{x}{2}\right)$$

- We recognize that $X \sim \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right) = \chi_1^2$

Example

- Let U be a uniform random variable on $[0, 1]$, and let $V = 1/U$. Find the density of V.



- Solution

1. Plot $v = 1/u$ and the pdf of U to find that V has possible values on $\{v : v \geq 1\}$.

2. Find $F_V(v)$. For $v \geq 1$,

$$F_V(v) = P(V \leq v) = P\left(\frac{1}{U} \leq v\right) = P\left(\frac{1}{v} \leq U \leq 1\right) = 1 - \frac{1}{v}$$

3. Find $f_V(v)$.

$$f_V(v) = \frac{1}{v^2}, \text{ for } v \geq 1.$$

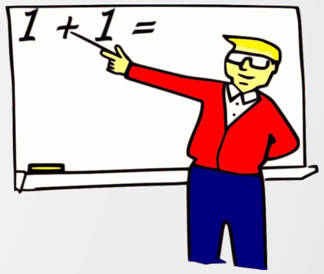
Proposition C. Let $Z = F(X)$. Then, $Z \sim \text{Unif}(0,1)$.

Sol. (Hint: The cdf of $\text{Unif}(0,1)$ is $F_Z(z) = z$ for $0 \leq z \leq 1$)

Follow the SOP:

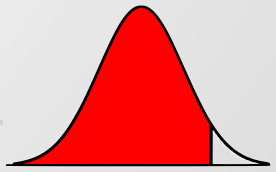
1. Plot $z = F(x)$. Note that cdf has to be non-decreasing and ranges from zero to one. Draw the pdf of X. (But we don't know what X is. That's okay!) $Z = F(X)$, so the support of Z is $z : 0 \leq z \leq 1$.

2. Find the cdf of Z. For $0 \leq z \leq 1$,



$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P\big(F(X) \leq z\big) \\ &= P\big(X \leq F^{-1}(z)\big) \\ &= F\big(F^{-1}(z)\big) = z \end{aligned}$$

3. From the cdf of Z, we recognize that $Z \sim U(0, 1)$.



Proposition D: The inverse method to generate RV. i

- Let U be a uniform on $[0, 1]$, and let $X = F^{-1}(U)$. Then the cdf of X is F.
- Proof.

$$F_X(x) = P(X \leq x) = P\left(F^{-1}(U) \leq x\right) = P\left(U \leq F(x)\right) = F(X)$$

- Reiteration: Suppose that X has a cdf, $F_X(x)$. Let $Y = F_X^{-1}(U)$, where $U \sim U(0, 1)$. Then, Y has the same distribution as X, i.e., $F_Y(y) = F_X(y)$.

Proposition D. ii

- Proof.

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) = P\left(F_X^{-1}(U) \leq y\right) \\
&= P\left(F_X\left(F_X^{-1}(U)\right) \leq F_X(y)\right) \\
&= P\left(U \leq F_X(y)\right) = F_X(y)
\end{aligned}
$$

Application of Proposition D. i

- If you would like to draw a sample from a distribution with cdf $F(\cdot)$. Then, set $Y = F^{-1}(U)$, $U \sim U(0,1)$. Y has cdf $F(\cdot)$.
- Example. How to generate random variables from an exponential distribution?
- Key idea: Suppose X has cdf F, then $F^{-1}(U)$ has the same distribution as X.

Application of Proposition D. ii

1. Pet $U \sim \mathrm{Unif}(0, 1)$.

2.$X \sim \mathrm{Exp}(\lambda)$, then X has cdf $F(x) = 1 - \exp(-\lambda x)$.

3. Find the inverse function of $F^{-1}(u)$. To find an x given y so that $F(x) = y$, we have

$$\begin{aligned} 1 - \exp(-\lambda x) = u \quad &\Rightarrow \quad 1 - u = \exp(-\lambda x) \\ &\Rightarrow \quad \ln(1 - u) = -\lambda x \\ &\Rightarrow \quad x = \frac{-\ln(1-u)}{\lambda} \end{aligned}$$

- Hence, we have

$$F^{-1}(u) = \frac{-\ln(1-u)}{\lambda}$$

Application of Proposition D. iii

4. By Proposition D, $\dfrac{-\ln(1-U)}{\lambda}$ has the distribution $\mathrm{Exp}(\lambda)$.

Application of Proposition D. iv

5. See computer experiments using python.

6. Proposition D shows that to generate a random variable with CDF $F(\,\cdot\,)$, if you have $F^{-1}(\,\cdot\,)$, you can generate it by $F^{-1}(U)$. Therefore, the care problem reduces to how to generate uniform random variable. This makes our life a lot easier.

Proposition B. i

- Let X be a continuous random variable with density $f(x)$ and let $Y = g(X)$ where g is differentiable, strictly monotonic function on some interval I. Suppose that $f(x) = 0$ if x is not in I. Then Y has the density function

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|$$

- for y such that $y = g(x)$ for some x, and $f_Y(y) = 0$ if $y \neq g(x)$ for any x in I. Here, g^{-1} is the inverse function of g; that is, $g^{-1}(y) = x$ if $y = g(x)$.

Proposition B. ii

An intuitive way (but not rigorous in math) to look at Proposition B: By change of variable in calculus, we have

$$
\begin{aligned}
P(X \in A) &= \int_A f_X(x)\,dx \\
&= \int_{A'} f_X\left(g^{-1}(y)\right)\left|\frac{dx}{dy}\right| dy \\
&= \int_{A'} f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right| dy \\
&= P(Y \in A')
\end{aligned}
$$

Therefore, we have $f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\dfrac{dg^{-1}(y)}{dy}\right|$.

Proof of Proposition B

[Case 1.] Suppose $g(\,\cdot\,)$ is strictly increasing. Then,

1. Find

$$F_Y(y) = P(Y \le y) = P\big(g(X) \le y\big) = P\big(X \le g^{-1}(y)\big) = F_X\big(g^{-1}(y)\big)$$

.

2. By differentiation, we obtain

$$f_Y(y) = \frac{d}{dy}F_Y(y) = f_X\big(g^{-1}(y)\big)\frac{d}{dy}g^{-1}(y)\,. \qquad (1)$$

Proof of Proposition B. Continued.

[Case 2.] Suppose $g(\,\cdot\,)$ is strictly decreasing. Then,

Find

$$F_Y(y) = P(Y \le y) = P\big(g(X) \le y\big) = P\big(X \ge G^{-1}(y)\big) = 1 - F_X\big(g^{-1}(y)\big)$$

By differentiation, we obtain

$$f_Y(y) = \frac{d}{dy}F_Y(y) = -f_X\big(g^{-1}(y)\big)\frac{d}{dy}g^{-1}(y)\,. \qquad (2)$$

Combining Eqs. (1) and (2), we obtain

$$F_Y(y) = f_X\big(g^{-1}(y)\big)\left|\frac{d}{dy}g^{-1}(y)\right|$$

Example 1

If $X \sim N(\mu, \sigma^2)$. Let $Y = a + bX$ for some constants a and b. Suppose $b \neq 0$. Find the distribution of Y.

Example 1 (cont.)

Solution

Let $y = g(x) = a + bx$. Then, we have

$$x = g^{-1}(y) = \frac{y-a}{b} \qquad \text{and} \qquad \frac{d}{dy} g^{-1}(y) = \frac{1}{b}$$

First, possible values of Y are in the set $\{y : y \in \mathbb{R}\}$.

Applying Proposition B, we obtain

Example 1 (cont.)

$$
\begin{aligned}
f_Y(y) &= f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right| \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\left(\frac{y-a}{b}-\mu\right)^2}{2\sigma^2}\right)\cdot\left|\frac{1}{b}\right| \\
&= \frac{1}{\sqrt{2\pi b^2\sigma^2}}\exp\left(-\frac{\left(y-(a+b\mu)\right)^2}{2b^2\sigma^2}\right)
\end{aligned}
$$

We recognize $Y \sim N\left(a+b\mu, b^2\sigma^2\right)$.

Example 2

- If $X \sim \mathrm{Exp}(\lambda)$. Let $Y = cX$ for some positive constant c. Find the distribution of Y.

Example 2 (cont.)

- Sol. Let $y = g(x) = cx$. Then,
- $x = g^{-1}(y) = \dfrac{y}{c}$ and $\dfrac{d}{dy}g^{-1}(y) = \dfrac{1}{c}$.
- Note that possible values of Y are in the set $\{y : y \geq 0\}$.

$$
\begin{aligned}
f_Y(y) &= f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right| \\
&= \lambda \exp\left(-\lambda\frac{y}{c}\right)\left|\frac{1}{c}\right| \\
&= \frac{\lambda}{c}\exp\left(-\frac{\lambda}{c}y\right), \text{ for } y \geq 0.
\end{aligned}
$$

- Hence, we recognize $\quad Y \sim \text{Exp}\left(\dfrac{\lambda}{c}\right)$

Example 3

- We say X has the Beta distribution with shape parameters $\alpha > 0$ and $\beta > 0$, denoted as $X \sim \mathrm{Beta}(\alpha, \beta)$, if it has the pdf

$$f(x) = \frac{1}{\mathrm{Beta}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

for $0 \leq x \leq 1$, where $\mathrm{Beta}(\alpha, \beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

Example 3 (cont.)

- A special case, $\text{Beta}(1, 1) \sim \text{Unif}(0, 1)$.
- In addition, we can calculate the following integral easily,

$$\int_0^1 x^4(1-x)^7\,dx = \text{Beta}(5, 8) = \frac{\Gamma(5)\Gamma(8)}{\Gamma(5+8)} = \frac{4!7!}{12!}$$