Special continuous random variables

January 14, 2021

X is an exponential RV with parameter λ if its pdf is

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

Denote $X \sim \mathcal{E}(\lambda)$

model arrival time of something, e.g. the amount of time until a piece of equipment breaks down, until a light bulb burns out, or until an accident occurs ...

$X \sim \mathcal{E}(\lambda)$ then

- $E(X) = \frac{1}{\lambda}$
- $\text{Var}(X) = \frac{1}{\lambda^2}$

The time until a small meteorite first lands anywhere in the Sahara desert is modeled as an exponential random variable with a mean of 10 days. The time is currently midnight. What is the probability that a meteorite first lands some time between 6am and 6pm of the first day?

- $X \sim \lambda$: time until a first lands
- $E(X) = 10 \Rightarrow \frac{1}{\lambda} = 10$ or $\lambda = \frac{1}{10}$
- pdf of X

$$f(x) = \begin{cases} \frac{1}{10}e^{-\frac{1}{10}x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

Solution

- $X \sim \lambda$: time until a first lands
- $E(X) = 10 \Rightarrow \frac{1}{\lambda} = 10$ or $\lambda = \frac{1}{10}$
- pdf of X

$$f(x) = \begin{cases} \frac{1}{10} e^{-\frac{1}{10}x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

- $X \sim \lambda$: time until a first lands
- $E(X) = 10 \Rightarrow \frac{1}{\lambda} = 10$ or $\lambda = \frac{1}{10}$
- pdf of X

$$f(x) = \begin{cases} \frac{1}{10}e^{-\frac{1}{10}x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

- between 6am to 6pm of the first day = $\{\frac{1}{4} \leq X \leq \frac{3}{4}\}$ (day)
-

$$P(\frac{1}{4} \leq X \leq \frac{3}{4}) = \int_{1/4}^{3/4} \frac{1}{10} e^{-\frac{1}{10}x} dx$$
$$\approx .0476$$

Continuous RV X is said to be normally distributed or Gaussian with parameter μ and σ^2 if its pdf is

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

for $-\infty < x < \infty$
Denote $X \sim \mathcal{N}(\mu, \sigma^2)$.

$X \sim \mathcal{N}(\mu, \sigma^2)$

- $E(X) = \mu$
- $\mathrm{Var}(X) = \sigma^2$

Bell shape, symmetric about the mean

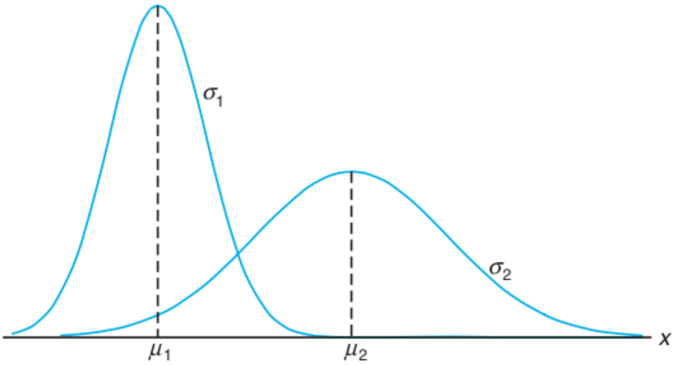


Figure: $\mu_1 < \mu_2, \sigma_1 < \sigma_2$

- Normal distibution is the most widely used distribution
- Many random phenomena obey a normal distribution
- Ex: the height and weight of a person, accuracy of shots from a gun... Link

Good approximation

- to approximate Binomial (n, p) when n is large
- *limiting distribution* of sample mean ...**broad base** for statistic inference (estimation and hypothesis testing), analysis of variance

- $Z \sim \mathcal{N}(0, 1)$ is standard normal distribution
- pdf

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

- cdf

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

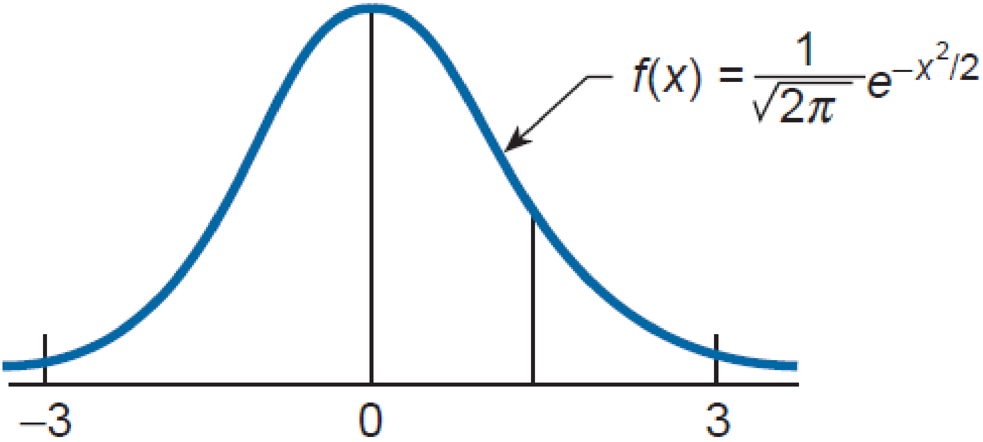


Figure: Pdf of Z

- Calculator
- Look up values in Normal Probability Table

Standard ormal probability table

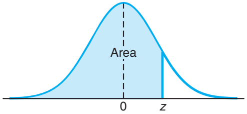


Table A.3 Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048

$P(Z \leq -2.54)$

Solution 1 - Calculator

1.

$$P(Z \leq -2.54) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{2.54} e^{-\frac{x^2}{2}} dx$$

$$= \lim_{a \to -\infty} \frac{1}{\sqrt{2\pi}} \int_{a}^{-2.54} e^{-\frac{x^2}{2}} dx$$

Substitute a by -10, -30, -50 ... and find the limit

Solution 2 - Look up the table value of normal probability

1. look up -2 in the first column
2. look up 54 in the first row
3. Intersection of the corresponding row and column

$$P(Z < -2.54) = .0055$$

Find

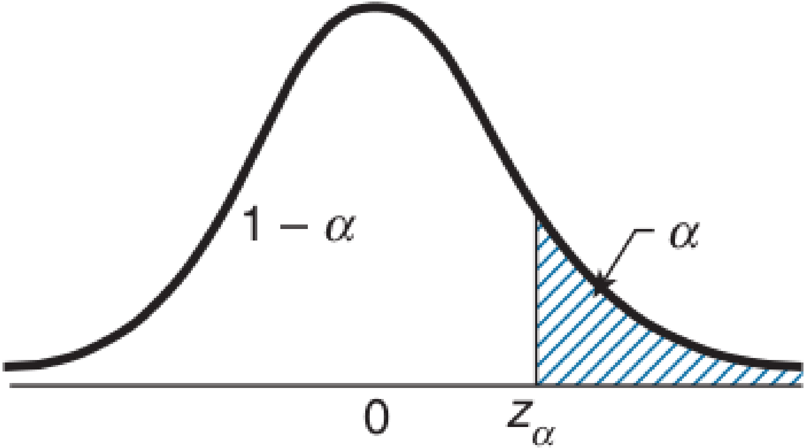
1. $P(Z > 2.33)$
2. $P(-1.65 < Z < 1.65)$
3. z such that $P(Z > z) = .95$

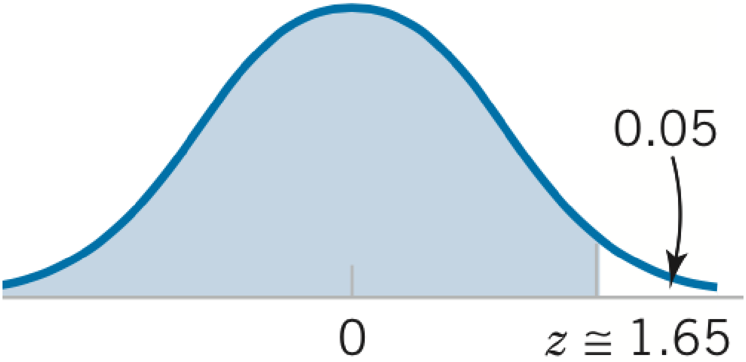
z_α is the $100(1-\alpha)$ percentile if

$$P(Z \leq z_\alpha) = 1 - \alpha$$

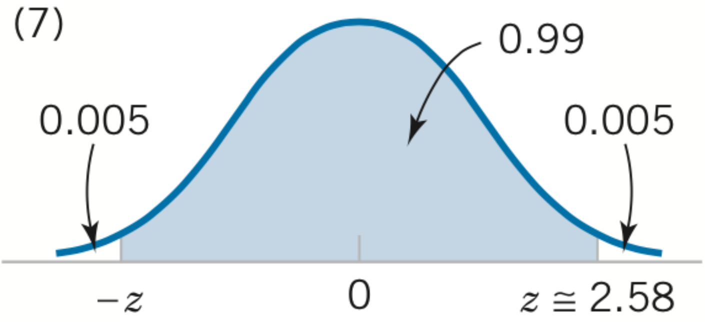
or

$$P(Z > z_\alpha) = \alpha$$


$1 - \alpha$
α
0
z_α



$$z_{.05} = 1.65$$



$$z_{.005} = 2.58$$

$$z_{.995} = -z_{.005} = -2.58$$

$$\boxed{z_{1-\alpha} = -z_\alpha}$$

because of the symmetry about y- axis of standard normal distribution

Normality is Preserved by Linear Transformations

If $X \sim \mathcal{N}(\mu, \sigma^2)$ then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

If $X \sim \mathcal{N}(\mu, \sigma^2)$ then

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

If $X \sim \mathcal{N}(\mu, \sigma^2)$ then

$$\begin{aligned} P(X \leq x) &= P(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}) \\ &= P(Z \leq \frac{x-\mu}{\sigma}) \end{aligned}$$

and

$$P(a \leq X \leq b) = P(Z < \frac{b-\mu}{\sigma}) - P(Z < \frac{a-\mu}{\sigma})$$

The annual snowfall at a particular geographic location is modeled as a normal random variable with a mean of $\mu = 60$ inches, and a standard deviation of $\sigma = 20$. What is the probability that this year's snowfall will be at least 80 inches?

Solution

- Snowfall $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = 60$, $\sigma = 20$

$$
\begin{aligned}
P(X \geq 80) &= 1 - P(X < 80) \\
&= 1 - P(Z < \frac{80 - \mu}{\sigma}) \\
&= 1 - P(Z < \frac{80 - 60}{20}) \\
&= 1 - P(Z < 1) = 1 - .8413 = .1687
\end{aligned}
$$

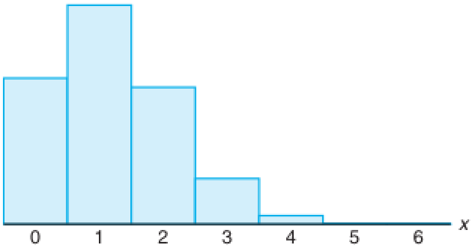


Figure 6.24: Histogram for $b(x; 6, 0.2)$.

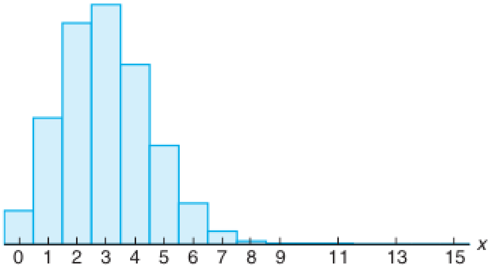


Figure 6.25: Histogram for $b(x; 15, 0.2)$.

- Suppose $Y \sim \text{ Bino}(n, p)$ where n is large and np is not too small
- Y can be approximated by $X \sim \mathcal{N}(np, np(1-p))$
- Y is discrete, X is continuous
- so we have to ”fill the gap”

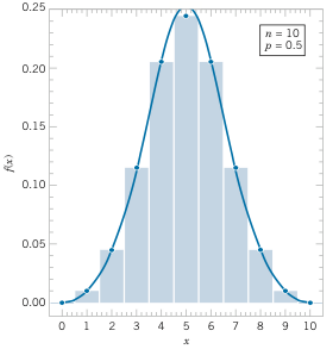


Figure 4-19 Normal approximation to the binomial distribution.

$$P(Y = i) \approx P(i - \frac{1}{2} < X < i + \frac{1}{2})$$

To approximate a binomial probability of $X \hookrightarrow Bin(n, p)$ with a normal distribution, a **continuity correction** is applied as follows:

$$P(X \leq x) = P(X \leq x + 0.5) \approx P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

The approximation is good for $np > 5$ *and* $n(1-p) > 5$.

Toss a fair coin 40 times. Y is the number of heads. Calculate $P(Y = 20)$ using normal approximation and direct computation.

Approximate Y by $X \sim \mathcal{N}(20, 10)$

$$
\begin{aligned}
P(Y = 20) &\approx P(19.5 < X < 20.5) \\
&= P\left(\frac{19.5 - 20}{\sqrt{10}} < Z < \frac{20.5 - 20}{\sqrt{10}}\right) \\
&= P(Z < .16) - P(Z < -.16) \\
&= .1272
\end{aligned}
$$

Exact value

$$P(Y = 20) = \binom{40}{20}(.5)^{40} = .1254$$

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

Solution

X: number of attending students $X \sim \text{Bino}(450, .3) \approx \mathcal{N}(135, 94.5)$

$$
\begin{aligned}
P(X > 150) &= P(X \geq 150.5) \\
&\approx P(Z \geq \frac{150.5 - 135}{\sqrt{94.5}}) \\
&= 1 - P(Z < 1.59) \\
&= .0559 \approx 5.6\%
\end{aligned}
$$