5.26 Normal Probabilities

Standard Normal Distribution

Examples

Sample Mean of Normal Observations

Definition: The Nor$(0,1)$ distrn is called the **standard normal** distribution.

Notation: The Nor$(0,1)$ is often denoted by Z.

The p.d.f. of the Nor$(0,1)$ is

$$\phi(z) \;\equiv\; \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad z \in \Re.$$

The c.d.f. is

$$\Phi(z) \;\equiv\; \int_{-\infty}^{z} \phi(t)\,dt, \quad z \in \Re.$$

Remarks:

$$\Pr(Z \leq a) \;=\; \Phi(a)$$

$$\Pr(Z \geq b) \;=\; 1 - \Phi(b)$$

$$\Pr(a \leq Z \leq b) \;=\; \Phi(b) - \Phi(a)$$

$$\Phi(0) \;=\; 1/2$$

$$\Phi(-b) \;=\; \Pr(Z \leq -b) \;=\; \Pr(Z \geq b) \;=\; 1 - \Phi(b)$$

$$\Pr(-b \leq Z \leq b) \;=\; \Phi(b) - \Phi(-b) \;=\; 2\Phi(b) - 1$$

Famous Nor$(0, 1)$ table values.

z	$\Phi(z) = \mathrm{Pr}(Z \leq z)$
0.00	0.5000
1.00	0.8413
1.28	$0.8997 \approx 0.90$
1.645	0.9500
1.96	0.9750
2.33	$0.9901 \approx 0.99$
3.00	0.9987
4.00	≈ 1.0000

Famous **Inverse** Nor$(0, 1)$ table values.

$\Phi^{-1}(p)$ is the value of z such that $\Phi(z) = p$.

p	$\Phi^{-1}(p)$
0.90	1.28
0.95	1.645
0.975	1.96
0.99	2.33
0.995	2.58

Example: $X \sim \text{Nor}(21, 4)$. Find $\Pr(19 < X < 22.5)$.

Standardizing, we get

$$
\begin{aligned}
&\Pr(19 < X < 22.5) \\
&= \Pr\left(\frac{19-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{22.5-\mu}{\sigma}\right) \\
&= \Pr\left(\frac{19-21}{2} < Z < \frac{22.5-21}{2}\right) \\
&= \Pr(-1 < Z < 0.75) \\
&= \Phi(0.75) - \Phi(-1) \\
&= \Phi(0.75) - [1 - \Phi(1)] \\
&= 0.7734 - [1 - 0.8413] \ = \ 0.6147.
\end{aligned}
$$

Example: Suppose that

Heights of men are $M \sim$ Nor$(68, 4)$ and

Heights of women are $W \sim$ Nor$(65, 1)$.

Select a man and woman *independently* at random.

Find the probability that the woman is taller than the man.

Note that

$$\begin{aligned} W - M &\sim \mathsf{Nor}(\mathsf{E}[W-M], \mathsf{Var}(W-M)) \\ &\sim \mathsf{Nor}(65-68, 1+4) \sim \mathsf{Nor}(-3, 5). \end{aligned}$$

Then

$$\begin{aligned} \Pr(W > M) &= \Pr(W - M > 0) \\ &= \Pr\Big(Z > \frac{0+3}{\sqrt{5}}\Big) \\ &= 1 - \Phi(3/\sqrt{5}) \\ &\approx 1 - 0.91 = 0.09. \end{aligned}$$

Sample Mean of Normal Observations

The sample mean of $X_1, \ldots, X_n$ is $\bar{X} \equiv \sum_{i=1}^{n} X_i/n$.

Corollary (of old Theorem): $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Nor}(\mu, \sigma^2) \Rightarrow \bar{X} \sim \text{Nor}(\mu, \sigma^2/n)$.

Proof: By previous work, as long as $X_1, \ldots, X_n$ are i.i.d. something, we have $\text{E}[\bar{X}] = \mu$ and $\text{Var}(\bar{X}) = \sigma^2/n$. Since $\bar{X}$ is a linear combination of independent normals, it's also **normal**. Done.

Remark: This result is *very significant!* As the number of observations increases, Var($\bar{X}$) gets *smaller* (while E[$\bar{X}$] remains constant).

In the upcoming statistics portion of the course, we'll learn that this makes the sample mean $\bar{X}$ an excellent **estimator** for the mean μ, which is typically unknown in practice.

Example: Suppose that $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim}$ Nor$(\mu, 16)$. Find the sample size n such that

$$\Pr(|\bar{X} - \mu| \leq 1) \ \geq \ 0.95.$$

How many observations should you take so that $\bar{X}$ will have a good chance of being close to μ?

Solution: Note that $\bar{X} \sim$ Nor$(\mu, 16/n)$. Then...

$$
\begin{aligned}
\Pr(|\bar{X}-\mu| \le 1) &= \Pr(-1 \le \bar{X}-\mu \le 1) \\
&= \Pr\left(\frac{-1}{4/\sqrt{n}} \le \frac{\bar{X}-\mu}{4/\sqrt{n}} \le \frac{1}{4/\sqrt{n}}\right) \\
&= \Pr\left(\frac{-\sqrt{n}}{4} \le Z \le \frac{\sqrt{n}}{4}\right) \\
&= 2\Phi(\sqrt{n}/4) - 1.
\end{aligned}
$$

Now we have to find n such that this probability is at least 0.95. . . .

$2\Phi(\sqrt{n}/4) - 1 \geq 0.95$ iff

$\Phi(\sqrt{n}/4) \geq 0.975$ iff

$$\frac{\sqrt{n}}{4} \geq \Phi^{-1}(0.975) = 1.96$$

iff $n \geq 61.47$ or 62.

So if you take the average of 62 observations, then $\bar{X}$ has a 95% chance of being within 1 of μ.