Normal Distribution

Let $Z$ be a standard Normal $Z \sim N(0, 1)$

$P(Z \leq z) = P(z)$ is given in Table A-2

Ex: $P(Z \leq 0) = .5$

$P(Z \leq -2.06) = .0197$

$P(Z \leq 0.32) = .6255$

What is $P(Z = 0)$?

The probability that a continuous RV takes any particular value is 0.

$P(Z = 0) = 0$ $P(Z < 0) = P(Z \leq 0)$

$P(Z > .32) = 1 - P(Z \leq .32) = 1 - .6255 = .3745$

$P(-.32 < Z < .32)$

$= P(Z < .32) - P(Z < -.32)$

$= .6255 - .3745$

$P(Z < -.32) = P(Z > .32)$

Can use the table in the other direction

What is the first quartile of a standard normal?

Find table entry closest to .25 and see which row and column it is

$P(Z < -.67) \approx .25$
If you don't have a standard normal, we standardize.
\[ Z = \frac{X - \mu}{\sigma} \]

where \( X \) is a normal with mean \( \mu \) and s.d. \( \sigma \)

Suppose the weight of a cookie is normally distributed with mean 11g and std. dev. 0.5g

\[
P(X < 11) = P\left(\frac{X - 11}{0.5} < \frac{11 - 11}{0.5}\right) = P(Z < 0) = 0.5
\]

\[
P(X < 10) = P\left(\frac{X - 11}{0.5} < \frac{10 - 11}{0.5}\right) = P(Z < -2) = 0.0228
\]

\[
P(X > 11.16) = P\left(Z > \frac{11.16 - 11}{0.5}\right) = P(Z > 0.32) = 1 - P(Z \leq 0.32)
\]

\[
= 1 - 0.6455 = 0.3745
\]

Suppose the cookie company wants to label their packages so that very few (1%) of the cookies weigh less than the listed weight. What should the labeled weight be?

\[
P(\text{cookie weight} < 10) = 0.01 \quad P(X < 10) = 0.01
\]

From standard normal table, find \( 0.01 \) .

\[
P(Z < -2.33) = 0.01
\]

\[
\Rightarrow Z = \frac{X - \mu}{\sigma} \quad \Rightarrow 0Z = X - \mu \quad \Rightarrow X = \mu + 0Z
\]

\[
= P(Z < -2.33) = P((1.5)Z + 11 < (1.5)(-2.33) + 11)
\]

\[
= P(X < 9.84)
\]

So 99% of the cookies will weigh at least 9.84 grams.

What is the number of chips such that 95% of the cookies will have at least that many chips?

Let \( Y \) = # chips in a cookie. Suppose \( Y \sim \text{Poisson} \) with mean 20.6

Normal is a good approximation for Poisson \( \mu = 20.6 \)

\[
P(Z < ?) = 0.05 \quad P(Z < -1.645) = 0.05
\]

\[
0.05 = P(\sigma Z + \mu < (1.645)(-1.645) + 20.6) = P(X < 13.13)
\]

95% of the cookies will have at least 13 chips.
What is the probability the average weight of a cookie in a bag of 50 will be at least 10.8 g?

\[
P(x \geq 10.8) = P \left( \frac{x - \mu_x}{\sigma_x} \geq \frac{10.8 - \mu_x}{\sigma_x} \right) = P \left( \frac{x - 11}{0.88} \geq \frac{10.8 - 11}{0.88} \right)
\]

\[
= P(z \geq -2.27) = 1 - P(z \leq -2.27) = 1 - .0116 = .9884
\]

Whereas \( P(x \geq 10.8) = P \left( \frac{x - 11}{5} \geq \frac{10.8 - 11}{5} \right) = P(z \geq -0.4) = 1 - P(z \leq -0.4) = 1 - .3446 = .6554 \)

Suppose the manufacturer wants to label their packages so that 99% of the bags will have an average cookie weight at least that large.

\[
P(z \leq -2.33) = .01
\]

\[
= P \left( \frac{z}{\sigma_x} \geq \frac{-2.33 (0.88)}{11} \right) = P(x \leq 10.79)
\]

99% of the bags will have a mean cookie weight of at least 10.79.

What if the manufacturer wants to claim that 95% of the bags will have an average weight in some interval? How do we find this interval? (Assume a symmetric interval)

\[
P(-z < z < z) = .95
\]

\[
P(z < z) = \frac{1 - .95}{2} = .025 = P(z < -1.96)
\]

So \( P(-1.96 < z < 1.96) = .95 \)

\[
P(-1.96 \sigma_x + M_x < x < 1.96 \sigma_x + M_x) = P(-1.96 \times 0.88 + 11 < x < 1.96 \times 0.88 + 11)
\]

\[
= P(10.83 < x < 11.17) = .95
\]

So 95% of the bags will have an average cookie weight between 10.83 g and 11.17 g.
Sampling Distributions

If we sample cookies, the number of the # chips in a cookie follows a Poisson distribution.

This is the sampling distribution of the number of chips.

What if we think about the average # chips per cookie in a bag?

Is this average the same for all bags?

No, it is random, and its distribution is derived from the sampling distribution of an individual cookie.

This is the sampling distribution of the mean.

Central Limit Theorem

If samples of size $n$ are drawn from a population with mean $\mu$ and std. dev. $\sigma$, then the sampling distribution of the sample means $\bar{x}$ will be approximately normally distributed with

mean $\mu_{\bar{x}} = \mu$ and std. dev. $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$\sigma_{\bar{x}}$ is called the (population) standard error (of the mean).

Note: If the population is normally distributed, then the sample mean is exactly normally dist'd. Otherwise, $\bar{x}$ is approximately normally dist'd, with the approximation getting better as $n$ increases. In general, $n > 30$ is good.

Ex.: Weight of a cookie is normally dist'd with $\mu = 11, \sigma = 0.5$

The mean weight of a cookie in a bag of 32 is normally distributed with mean $\mu_{\bar{x}} = \mu = 11$ and std. dev. $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.5}{\sqrt{32}} = 0.088g$

Notice as $n \rightarrow \infty, \sigma_{\bar{x}} \rightarrow 0$