Summary on normal distribution

1. The Normal Distribution

(a) The normal distribution with mean μ and variance σ^2: $X \sim N(\mu, \sigma^2)$

(b) The standard normal distribution: $\frac{X-\mu}{\sigma} = Z \sim N(0,1)$

  i. The distribution curve is bell-shape.

  ii. Use the standard normal table of $P(Z < z)$

  iii. Z is symmetric about 0: $P(Z < z) = 1 - P(Z < -z)$

  iv. $P(Z < -z) = P(Z > z)$

  v. Standard normal percentiles and critical values:

Percentile	90	95	97.5	99	99.5
α (tail area)	.1	.05	.025	.01	.005
z_α	1.28	1.645	1.96	2.33	2.58

  vi. Values of $z_{\alpha/2}$:

$1-\alpha$	.80	.90	.95	.98	.99
$z_{\alpha/2}$	1.28	1.645	1.96	2.33	2.58

2. The Normal Approximation to the Binomial Distribution

If $X \sim \text{Bin}(n,p)$ with large n, then $Z = \frac{X-np}{\sqrt{np(1-p)}}$ is approximately standard normal.

(a) Type of the Problems

A discrete random variable X is involved. $X \sim \text{Bin}(n,p)$. Probability is required. Typically n is large, $np \geq 5$, $n(1-p) \geq 5$: approximation is appropriate.

(b) Method of approximation

Step 1 $\mu = np$, $\sigma^2 = np(1-p)$
Step 2 Use $N(\mu, \sigma^2) = N(np, np(1-p))$
Step 3 Remember to use continuity correction.

  i. $P(X = 3) = P(\frac{2.5-\mu}{\sigma} \leq Z \leq \frac{3.5-\mu}{\sigma})$

  ii. $P(X < 3) = P(Z \leq \frac{2.5-\mu}{\sigma})$

  iii. $P(X \leq 3) = P(Z \leq \frac{3.5-\mu}{\sigma})$

  iv. $P(X > 3) = P(Z \geq \frac{3.5-\mu}{\sigma})$

  v. $P(X \geq 3) = P(Z \geq \frac{2.5-\mu}{\sigma})$

Example: $P(0 < X \leq 3) = P(\frac{.5-\mu}{\sigma} \leq Z \leq \frac{3.5-\mu}{\sigma})$

3. **Sampling Distributions**

(a) The r.v.'s $X_1, X_2, \ldots, X_n$ are a random sample (iid) of size n if

i. $X_1, X_2, \ldots, X_n$ are independent.

ii. Every X_i has the same probability distribution.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with mean μ and standard deviation σ, then

i. $\mathrm{E}(\overline{X}) = \mu$

ii. $\mathrm{Var}(\overline{X}) = \sigma^2/n$ and $\sigma_{\overline{X}} = \sigma/\sqrt{n}$

(c) Let $X_1, X_2, \ldots, X_n$ be a random sample from any distribution with mean μ and variance σ^2. Then for large n ($n \geq 30$),

i. $\overline{X}$ is approximately $N(\mu, \sigma^2/n)$

ii. Central Limit Theorem (CLT):

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \quad \text{is approximately} \quad N(0,1)$$

(d) For unknown σ

If $X_1, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$, and $S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/(n-1)$, then

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Example

Review on Sampling Distribution of the Mean

i. Type of the Problems

(Population) mean μ and standard deviation σ of a random variable X is given. Then it asks for probability concerning the sample mean $\overline{X}$. (In some problems, probability of sum is required).

ii. II. Method

A. If $X \sim N(\mu, \sigma^2)$, $\overline{X} \sim N(\mu, \sigma^2/n)$, $n \geq 1$

B. If X is not normal, $\overline{X} \simeq N(\mu, \sigma^2/n)$, $n \geq 30$

iii. In any case, we calculate probability using $N(\mu, \sigma^2/n)$.
Do not use continuity correction.
Eg. X has mean $\mu = 5$, standard deviation $\sigma = 20$, $n = 100$.
$\overline{X}$ is approximately $N(\mu, \sigma^2/n) = N(5, 4)$.
Hence $P(\overline{X} > 6) = P(Z > \frac{6-5}{2}) = P(Z > .5) = .3085$.