Math 3501 - Probability and Statistics I

5.3 - Several random variables

generalizes some of notions from Chp 4 to any number of r.v.s.

Linear combination independent random variables

Theorem

Let $X_1, X_2, \ldots, X_n$ be n independent random variables with respective means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$.

The mean and variance of

$$Y = \sum_{i=1}^{n} a_i X_i \ ,$$

where $a_1, a_2, \ldots, a_n$ are real constants, are, respectively,

expectation is linear

always holds $\longrightarrow$

$$\mu_Y = \sum_{i=1}^{n} a_i \mu_i$$

$\rightarrow E[a_1 X_1 + \ldots + a_m X_m] = a_1 E[X_1] + \cdots + a_m E[X_m]$

and

when independency is really needed! $\longrightarrow$

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2 \ .$$

$\rightarrow \mathrm{Var}(a_1 X_1 + \cdots + a_m X_m) = a_1^2 \mathrm{Var}(X_1) + \cdots + a_m^2 \mathrm{Var}(X_m)$

Statistic

Definition

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n.

Any function of the sample observations $X_1, X_2, \ldots, X_n$ that does not have any unknown parameters is called a *statistic*.

$$\left.\begin{array}{l} Y = \sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n \\ Y = \max\{x_1, x_2, \ldots, x_n\} \end{array}\right\} \text{statistic}$$

$Y = \sum_{i=1}^{n} (x_i - \mu)^2$ IS NOT STATISTIC if μ is not a STATISTIC

Sample mean and sample variance

Definition

main topic of today's discussion

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n.

Its *sample mean* is

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

and its *sample variance* is

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 .$$

Example

The sample mean and the sample variance of a random sample of size n are statistics.

Mean and variance of the sample mean

σ^2 = variance of each X_i

Theorem

i.i.d

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution with mean μ and variance σ^2.

The mean and variance of the sample mean $\bar{X}$ are

$$\mu_{\bar{X}} = E\left(\bar{X}\right) = \mu$$

and

Variance of $\bar{X}$

$$\sigma^2_{\bar{X}} = \text{Var}\left(\bar{X}\right) = \boxed{\frac{\sigma^2}{n}}.$$

as n increases our estimate $\bar{X}$ for μ improves in quality as measured by $\sigma^2_{\bar{X}}$

$$\mu_{\bar{X}} = E[\bar{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[X_i]}_{\mu} = \frac{1}{n}\cdot n\mu = \mu$$

$$\sigma^2_{\bar{X}} = \text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\overbrace{\text{Var}(X_i)}^{\sigma^2} = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

$X_1, \ldots, X_n$ independent

Math 3501 - Probability and Statistics I

5.4 - The moment generating function technique

helps us determine the distribution of sums of random variables

Linear combination of independent random variables mgf

Theorem

Let $X_1, X_2, \dots, X_n$ be independent random variables with respective moment-generating functions $M_{X_i}(t)$ defined on intervals of the form $(-h_i, h_i)$, for some positive constants h_i, $i = 1, 2, \dots, n$.

The moment-generating function of the linear combination

$$Y = \sum_{i=1}^{n} a_i X_i$$

is given by

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t) \ ,$$

where t satisfies each one of the conditions $-h_i < a_i t < h_i$ for $i = 1, 2, \dots, n$.

mgf of the linear combination of $X_1, \dots, X_m$ is the product of mgf of $X_1, \dots, X_m$

special cases:

$Y = \sum_{i=1}^{m} X_i \Rightarrow a_1 = a_2 = \dots = a_m = 1$

$Y = \frac{1}{m} \sum_{i=1}^{m} X_i \Rightarrow a_1 = a_2 = \dots = a_m = \frac{1}{m}$

all independent

$$M_Y(t) = E[e^{tY}] = E\left[e^{t\sum_{i=1}^{m} a_i X_i}\right] = E\left[e^{a_1 t X_1 + a_2 t X_2 + \dots + a_m t X_m}\right] = E\left[e^{a_1 t X_1} \cdot e^{a_2 t X_2} \cdots e^{a_m t X_m}\right]$$

$$= E\left[e^{a_1 t X_1}\right] \cdot E\left[e^{a_2 t X_2}\right] \cdots E\left[e^{a_m t X_m}\right] = M_{X_1}(a_1 t) \cdot M_{X_2}(a_2 t) \cdots M_{X_m}(a_m t)$$

Sample mean mgf

iid ⇒ identically distributed r.v.s have same m.g.f

Corollary

If $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution with mgf $M(t)$, where $-h < t < h$.

(a) The mgf of the sum $Y = \sum_{i=1}^{n} X_i$ is

$a_1 = a_2 = \cdots = a_n = 1$

$$M_Y(t) = \prod_{i=1}^{n} M(t) = [M(t)]^n , \qquad -h < t < h .$$

(b) The mgf of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is

$a_1 = a_2 = \cdots = a_n = \frac{1}{n}$

$$M_{\bar{X}}(t) = \prod_{i=1}^{n} M\left(\frac{t}{n}\right) = \left[M\left(\frac{t}{n}\right)\right]^n , \qquad -h < \frac{t}{n} < h .$$

Example

Let $X_1, X_2, \ldots, X_n$ denote the outcomes of n Bernoulli trials, each with probability of success p.

Find the mgf of $Y = \sum_{i=1}^{n} X_i$. $\Longrightarrow$ $Y \sim \text{Binomial}(n, p)$

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli (p) distribution

Their common mgf is $M(t) = 1 - p + pe^t$, $t \in \mathbb{R}$.

$$M_Y(t) = E\left[e^{tY}\right] = E\left[e^{t\sum X_i}\right] = E\left[e^{tX_1 + tX_2 + \cdots + tX_n}\right] =$$

$$= E\left[e^{tX_1}\right] \cdot E\left[e^{tX_2}\right] \cdots E\left[e^{tX_n}\right]$$

$$= \underbrace{M(t) \cdot M(t) \cdots M(t)}_{n \text{ times}} = \left(M(t)\right)^n = \underbrace{\left(1 - p + pe^t\right)^n}_{\text{mgf of binomial}(n,p)}$$

Example

Let $X_1, X_2, \ldots, X_n$ be the observations of a random sample of size n [redacted] from the exponential distribution having mean θ.

Find the mgf of $Y = \sum_{i=1}^{n} X_i$.

Let $X_1, \ldots, X_n \sim$ Exponential (θ).

Then the m.g.f of $X_1, X_2, \ldots, X_n$ is $M(t) = \frac{1}{1-\theta t}$, $t < \frac{1}{\theta}$

and so

$$M_Y(t) = \left(M(t)\right)^n = \left(\frac{1}{1-\theta t}\right)^n = \frac{1}{(1-\theta t)^n}, \quad t < \frac{1}{\theta}$$

$\Rightarrow Y \sim$ Gamma (n, θ)

Math 3501 - Probability and Statistics I

5.5 - Random functions associated with the normal distribution

Linear combination of normally distributed r.v.

Theorem

Let $X_1, X_2, \ldots, X_n$ be n independent normally distributed random variables with respective means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$.

The linear combination

$$Y = \sum_{i=1}^{n} c_i X_i$$

has the normal distribution

$$N\left(\underbrace{\sum_{i=1}^{n} c_i \mu_i}_{E[Y]}, \underbrace{\sum_{i=1}^{n} c_i^2 \sigma_i^2}_{Var(Y)}\right).$$

proof : Use the mgf technique from Sec. 5.4 (Exercise)

Very special feature of the normal distribution

Example

Let X_1 and X_2 be independent normally distributed random variables with respective distributions $N(5,2)$ and $N(2,1)$.

Find $P(X_1 > X_2)$.

Let $X_1 \sim N(5,2)$ and $X_2 \sim N(2,1)$ be independent r.v.

$$P(X_1 > X_2) = P(\underbrace{X_1 - X_2}_{\text{linear combination of } X_1 \text{ and } X_2} > 0) = ???$$

Let $Y = X_1 - X_2 \sim N(\underset{E[Y]}{3}, \underset{Var(Y)}{3})$

$E[Y] = E[X_1 - X_2] = E[X_1] - E[X_2] = 5 - 2 = 3$

$Var(Y) = Var(X_1 - X_2) \underset{\text{X}_1 \text{ and } X_2 \text{ independent}}{=} (1)^2 \cdot Var(X_1) + (-1)^2 \cdot Var(X_2) = Var(X_1) + Var(X_2) = 2 + 1 = 3$

$$P(X_1 > X_2) = P(X_1 - X_2 > 0)$$

$$= P(Y > 0) = P\Bigg(\overbrace{\frac{Y-3}{\sqrt{3}}}^{Z} > \frac{0-3}{\sqrt{3}}\Bigg)$$

$Y \sim N(3,3)$

$\Downarrow$

$Z = \dfrac{Y - \overset{\mu}{3}}{\underset{\sigma}{\sqrt{3}}} \sim N(0,1)$

use the following result from Sec 3.3

If $X \sim N(\mu, \sigma^2) \Rightarrow Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$

$$= P\left(Z > -\frac{3}{\sqrt{3}}\right)$$

$$= P(Z > -\sqrt{3})$$

$$= 1 - P(Z < -\sqrt{3})$$

$$= 1 - \phi(-\sqrt{3})$$

$$= 1 - (1 - \phi(\sqrt{3}))$$

$$= \phi(\sqrt{3}) \approx \phi(1.73) \underset{\text{table}}{\approx} 0.9582$$

Sample mean of normally distributed r.v.

Corollary

Let $X_1, X_2, \ldots, X_n$ be observations of a random sample of size n from the normal distribution $N(\mu, \sigma^2)$.

The distribution of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ *is* $N\left(\mu, \frac{\sigma^2}{n}\right)$

$E[\bar{x}]$ Var$(\bar{x})$

Example

Let $X_1, X_2, \ldots, X_{64}$ be a $\overbrace{\text{random sample}}^{iid}$ from the $N(\underset{\uparrow \mu=50}{50}, \underset{\uparrow \sigma^2=16}{16})$ distribution.

Find $P(49 < \bar{X} < 51)$.

Since $X_1, \ldots, X_{64} \sim N(50, 16)$, then

$$\boxed{\bar{X} \sim N\left(\underset{\uparrow \mu}{50}, \underset{\uparrow \sigma^2/n}{\frac{16}{64}}\right)} \leftarrow \text{From the previous corollary}$$

$$Z = \frac{\bar{X} - 50}{\sqrt{\frac{16}{64}}} = \frac{\bar{X} - 50}{1/2} \sim N(0,1)$$

using that: $X \sim N(\mu, \sigma^2) \Rightarrow Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$

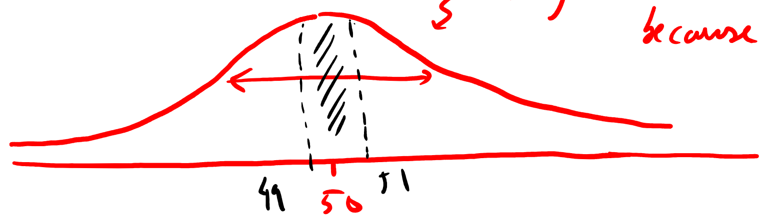
$$P(49 < \bar{X} < 51) = P\left(\frac{49-50}{1/2} < \underbrace{\frac{\bar{X}-50}{1/2}}_{Z \sim N(0,1)} < \frac{51-50}{1/2} \right)$$

$$= P(-2 < Z < 2)$$

$$= \phi(2) - \underbrace{\phi(-2)}_{1-\phi(2)} = \phi(2) - \overbrace{(1-\phi(2))}^{\phi(-2)}$$

$$= 2\phi(2) - 1 = \underbrace{2(0.9772)}_{\text{table}} - 1 = \underbrace{\dots}_{\text{very close to } 1}$$

$X_1, X_2, \dots, X_{64} \longrightarrow$ pdf $N(50, 16)$



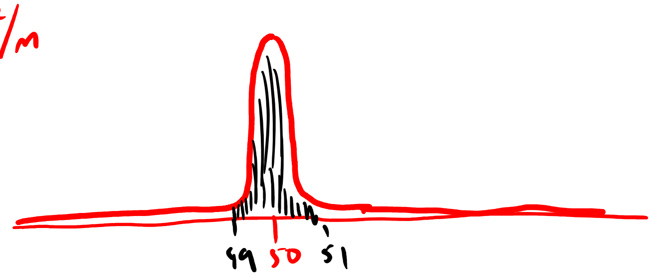
$\leftarrow$ very wide bell curve because $\sigma^2 = 16$ is large

If we tried to compute $P(49 < X_i < 51)$ = area shaded in black = small number

$$\bar{X} = \frac{1}{m} \sum_{i=1}^{64} X_i \sim N\left(50, \frac{1}{4}\right) \longrightarrow \text{pdf looks like}$$

$\uparrow$
σ^2/m

$P(49 < \bar{X} < 51) =$ nearly 1



Math 3501 - Probability and Statistics I

5.6 - $\underbrace{\text{Central Limit Theorem}}_{}$

The most fundamental result to Probability and Statistics

Central Limit Theorem

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Theorem

i.i.d

Let $\bar{X}$ be the mean of a random sample $X_1, X_2, \ldots, X_n$ of size n from a distribution with a finite mean μ and a finite positive variance σ^2.

No requirements concerning the distribution of $X_1, X_2, \ldots, X_n$

Then the distribution of

$$\boxed{W = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}} = \frac{\overbrace{\sum_{i=1}^{n} X_i}^{Y} - n\mu}{\sqrt{n}\sigma}$$

approaches a standard normal distribution $N(0,1)$ in the limit as $n \to \infty$.

Consequence: For sufficiently large n:

1) $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is approximately $N\left(\mu, \frac{\sigma^2}{n}\right)$ distributed

2) $Y = \sum_{i=1}^{n} X_i$ is approximately $N\left(n\mu, n\sigma^2\right)$ distributed

Notes: The Central Limit Theorem describes the limit behavior of the distribution of the sequence of random variables

$$W = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}, \qquad n = 1, 2, 3, \ldots$$

Note how powerful the statement is: regardless of the distribution from which we are sampling (provided it has finite mean μ and finite variance σ^2), the limiting distribution of W is $N(0, 1)$.

For sufficiently large n, the Central Limit Theorem may be used to approximate the cdf of W:

$$P(W \le w) \approx \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \Phi(w)$$

In general the approximation is reasonably good for $n \ge 30$.

- In the special case of symmetric and unimodal distributions of the continuous type, $n \ge 5$ may be enough for an adequate approximation.

Example

Let $\bar{X}$ be the mean of a random sample of $n = 25$ currents (in milliamperes) in a strip of wire in which each measurement has a mean of 15 and a variance of 4.

Find $P(14.4 < \bar{X} < 15.6)$.

$X_1, X_2, \ldots, X_{25}$ random sample from a distribution with mean $\mu = 15$ and variance $\sigma^2 = 4$

Using the CLT we know that

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - 15}{2/\sqrt{25}} = \frac{\bar{X} - 15}{2/5} \text{ is approx } N(0,1)$$

Then

$$P(14.4 < \bar{X} < 15.6) = P\left(\frac{14.4 - 15}{2/5} < \underbrace{\frac{\bar{X} - 15}{2/5}}_{Z} < \frac{15.6 - 15}{2/5}\right)$$

$$= P(-1.5 < Z < 1.5) \underset{Z \text{ is approx } N(0,1)}{\approx} \phi(1.5) - \phi(-1.5) = \phi(1.5) - (1 - \phi(1.5))$$

$$= 2\phi(1.5) - 1 \approx 2(0.9332) - 1 \approx \ldots$$

Example

Let $X_1, X_2, \ldots, X_{20}$ denote a random sample of size 20 from the uniform distribution $U(0,1)$.

If $Y = \sum_{i=1}^{20} X_i$, find $P(8.5 \leq Y \leq 11.7)$.

Since $X_1, X_2, \ldots, X_{20} \sim \text{Uniform}(0,1) \Rightarrow \left\{ \begin{array}{l} \mu = E[X_i] = \frac{1}{2} \\ \sigma^2 = Var(X_i) = \frac{1}{12} \end{array} \right\}$ check the formula tables

CLT $\Rightarrow$ $Z = \dfrac{Y - n\mu}{\sqrt{n}\,\sigma} = \dfrac{Y - 20 \cdot \frac{1}{2}}{\sqrt{20} \cdot \sqrt{\frac{1}{12}}} = \dfrac{Y - 10}{\sqrt{5/3}}$ is approx $N(0,1)$

Thus $P(8.5 < Y < 11.7) = P\left(\dfrac{8.5 - 10}{\sqrt{5/3}} < \underbrace{\dfrac{Y - 10}{\sqrt{5/3}}}_{Z} < \dfrac{11.7 - 10}{\sqrt{5/3}} \right)$

$$= P\left(\frac{-1.5}{\sqrt{5/3}} < Z < \frac{1.7}{\sqrt{5/3}} \right) \approx \phi\left(\frac{1.7}{\sqrt{5/3}}\right) - \phi\left(\frac{-1.5}{\sqrt{5/3}}\right) = \cdots\cdots$$

CLT approx

use table to complete evaluation!

Math 3501 - Probability and Statistics I

half-unit correction!

5.7 - $\underbrace{\text{Approximations for discrete distributions}}$

how to apply CLT to discrete r.v.s whose support is a subset of $\mathbb{N}$

ISSUE

if X is discrete:

$P(X=k)$ is a positive number if k is in the support of X

For a continuous r.v $\quad P(X=k)=0$

If X is discrete with support contained in $\mathbb{N}$, and we use the CLT to approx $P(X=k)$ we cannot just do $P\underbrace{(Z = \text{some number})}_{\text{would be zero!}}$

CLT approximation for the binomial distribution

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli distribution with mean $\mu = p$ and variance $\sigma^2 = p(1-p)$, where $0 < p < 1$.

Then $Y = \sum_{i=1}^{n} X_i$ is $b(n, p)$. [we've seen this today !!]

The central limit theorem states that the distribution of

$$W = \frac{Y - np}{\sqrt{np(1-p)}} = \frac{\bar{X} - \overset{\mu}{p}}{\sqrt{\underbrace{p(1-p)}_{\sigma^2}/n}}$$

is $N(0, 1)$ in the limit as $n \to \infty$.

For n sufficiently large, the distribution of Y is approximately $N[np, np(1-p)]$:

- probabilities for the binomial distribution $b(n, p)$ can be approximated using the normal distribution.
- Rule of thumb: n is sufficiently large if $np \geq 5$ and $n(1-p) \geq 5$

Half-unit correction for continuity

If V is $N\left(\mu, \sigma^2\right)$, then $P(a < V < b)$ equals the area bounded by the pdf of V, the horizontal-axis and the vertical lines $v = a$ and $v = b$.

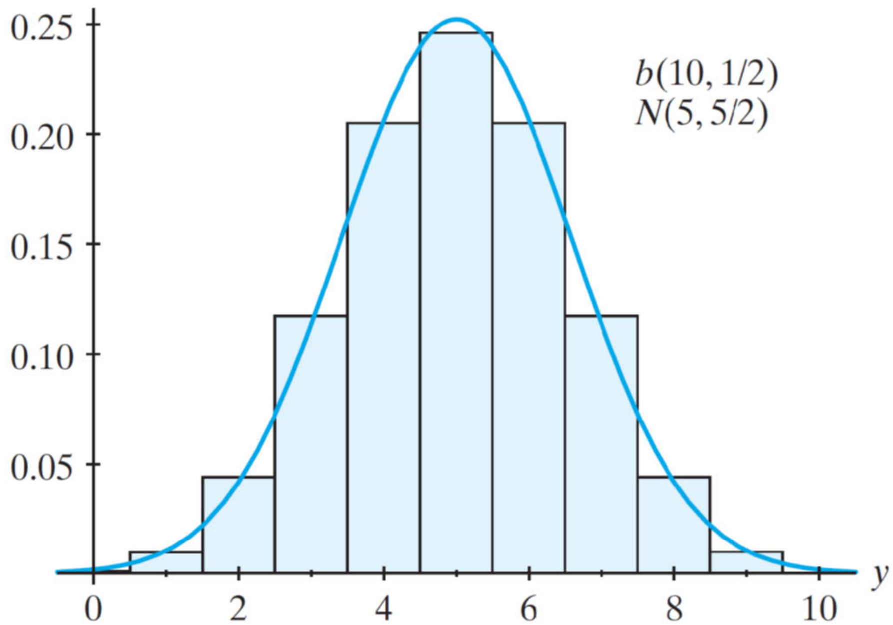
If Y is $b(n,p)$, then $P(Y = k)$ can be represented by the area of a rectangle with a height of $P(Y = k)$ and a base of length 1 centered at k

For an integer k

$$
\begin{aligned}
P(Y=k) &= \overbrace{P(k-1/2 < Y < k+1/2)}^{k \text{ is the only integer in } (k-1/2,\, k+1/2)} \\
&= P\left(\frac{k-1/2-np}{\sqrt{npq}} < \frac{Y-np}{\sqrt{npq}} < \frac{k+1/2-np}{\sqrt{npq}}\right) \\
&\approx \Phi\left(\frac{k+1/2-np}{\sqrt{npq}}\right) - \Phi\left(\frac{k-1/2-np}{\sqrt{npq}}\right)
\end{aligned}
$$

CLT approx

$$\underbrace{Y=k}_{k-1/2 < Y < k+1/2}$$

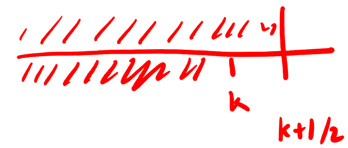


0.25
0.20
0.15
0.10
0.05
0
2
4
6
8
10
y
b(10, 1/2)
N(5, 5/2)

In general, if Y is $b(n,p)$, n is sufficiently large, and $k = 0, 1, \ldots, n$, then

include k

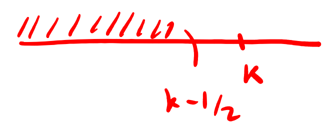
$$P(Y \leq k) \approx \Phi\left(\frac{k + 1/2 - np}{\sqrt{npq}}\right)$$



and

exclude

$$P(Y < k) \approx \Phi\left(\frac{k - 1/2 - np}{\sqrt{npq}}\right)$$

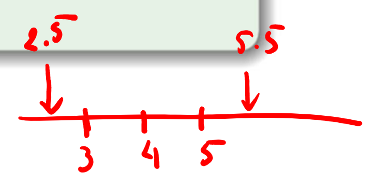


because in the first case k is included and in the second it is not.

Example

Suppose Y is $b(10, 1/2)$.

Use the normal distribution to approximate $P(3 \leq Y < 6)$.



$Y \sim b(10, 1/2)$ $\quad n = 10, \; p = 1/2$

CLT $\Rightarrow$ $Z = \dfrac{Y - np}{\sqrt{np(1-p)}} = \dfrac{Y - 10 \cdot 1/2}{\sqrt{10 \cdot \frac{1}{2} \cdot \frac{1}{2}}} = \dfrac{Y - 5}{\sqrt{5/2}}$ is approx $N(0,1)$

$$P(\underset{\text{include 3}}{3 \leq} Y \underset{\text{exclude 6}}{< 6}) \underset{\text{half unit correction}}{=} P(\underbrace{2.5 \leq Y \leq 5.5}_{\text{3, 4, 5 are the only integers in the interval } (2.5, 5.5)}) =$$

$$= P\left(\frac{2.5 - 5}{\sqrt{5/2}} \leq \underbrace{\frac{Y - 5}{\sqrt{5/2}}}_{Z} \leq \frac{5.5 - 5}{\sqrt{5/2}} \right)$$

$$= P\left(-\frac{2.5}{\sqrt{5/2}} \leq Z \leq \frac{0.5}{\sqrt{5/2}} \right) \underset{\substack{\uparrow \\ \text{CLT} \\ \text{approx}}}{\approx} \phi\left(\frac{0.5}{\sqrt{5/2}}\right) - \phi\left(-\frac{2.5}{\sqrt{5/2}}\right)$$

$\approx$ use table etc....

CLT approximation for the Poisson distribution

Let $X_1, \dots, X_n$ be a random sample from a Poisson distribution with mean $\lambda = 1$.

Then $Y = \sum_{i=1}^{n} X_i$ is Poisson distributed with mean $\lambda = n$ (and variance $\lambda = n$).

The central limit theorem states that the distribution of

$\mu = \lambda = 1$

$n\mu$

$$W = \frac{Y - n}{\sqrt{n}}$$

$\sqrt{n} \cdot \sigma \quad \sigma = \sqrt{\lambda} = 1$

is $N(0, 1)$ in the limit as $n \to \infty$.

In general, for sufficiently large λ, the distribution of

$$W = \frac{Y - \lambda}{\sqrt{\lambda}}$$

approx $N(0,1)$

is approximately $N(0, 1)$.

The approximation is reasonably good provided $\lambda \geq 10$ and an appropriate continuity correction is used:

$$\begin{aligned} P(Y = k) &= P(k - 1/2 < Y < k + 1/2) \\ &= P\left(\frac{k - 1/2 - \lambda}{\sqrt{\lambda}} < \frac{Y - \lambda}{\sqrt{\lambda}} < \frac{k + 1/2 - \lambda}{\sqrt{\lambda}}\right) \\ &\approx \Phi\left(\frac{k + 1/2 - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{k - 1/2 - \lambda}{\sqrt{\lambda}}\right) \end{aligned}$$

same as done for binomial!!

for any nonnegative integer k.

Example

Suppose Y is Poisson with mean 20.

Use the normal distribution to approximate $P(16 < Y \leq 21)$.

$Y \sim \text{Poisson}(20) \qquad \lambda = 20$

CLT $\Rightarrow$ $Z = \dfrac{Y - \lambda}{\sqrt{\lambda}} = \dfrac{Y - 20}{\sqrt{20}}$ is approx $N(0,1)$

$P(\underbrace{16 < Y \leq 21}_{Y = 17, 18, 19, 20, 21}) \underset{\text{half unit correction}}{=} P(16.5 \leq Y \leq 21.5) = \longrightarrow$

$$P\left(\frac{16.5 - 20}{\sqrt{20}} \le \underbrace{\frac{Y - 20}{\sqrt{20}}}_{Z} \le \frac{21.5 - 20}{\sqrt{20}} \right)$$

$$= P\left(-\frac{3.5}{\sqrt{20}} \le Z \le \frac{1.5}{\sqrt{20}} \right)$$

$$\underset{\text{CLT approx.}}{\approx} \quad \phi\left(\frac{1.5}{\sqrt{20}}\right) - \phi\left(-\frac{3.5}{\sqrt{20}}\right) = \underset{\text{use table}}{\cdots\cdots}$$