Chapter 5 Distributions of Functions of random variables

Section 5.7 Approximations for discrete distributions

➢ Motivation: CLT applies to discrete distributions as well. In this section, we illustrate how the normal distribution can be used to approximate probabilities for discrete distributions.

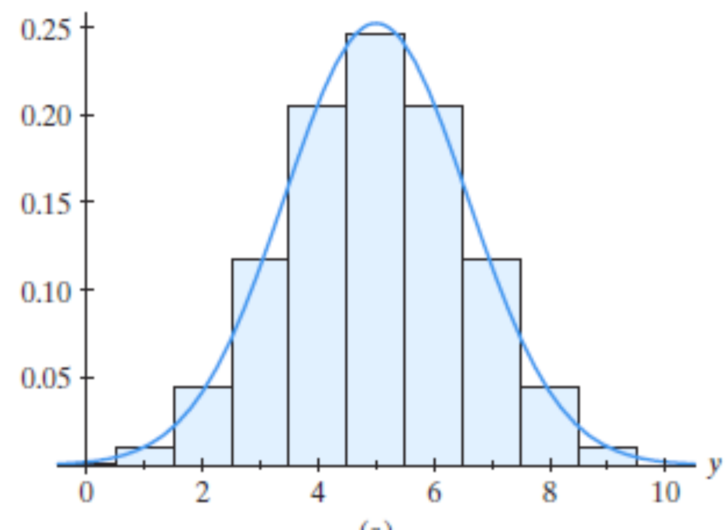
- Histogram for discrete distribution

Consider a discrete distribution with pmf $f(x): \overline{S} \to [0,1]$ with $\overline{S} = \{0,1,\ldots\}$.

Then the histogram for the discrete distribution is $h(x) = f(k),\ x \in \left(k-\frac{1}{2}, k+\frac{1}{2}\right)$.

$$k = 0,1,\ldots,n.$$

Then $P(X = k)$ is the area of the rectangle with a height of $P(X = k)$ and a base of length 1 centered at k.



➢ Half with correction for continuity.

- When using CLT and normal distribution to approximate probabilities for discrete distributions, we have

$$P(X = k) = P(k - \frac{1}{2} < X < k + \frac{1}{2})$$

pmf $f(x)$ [hard to derive]

Approximated by the normal distribution according to CLT [easier to derive]

➢ Binomial distribution

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from Bernoulli distribution $b(1, p)$,

whose mean is p and variance $p(1-p)$. Then $Y = \sum_{i=1}^{n} X_i \sim b(n, p)$ with mean $\mu = np$

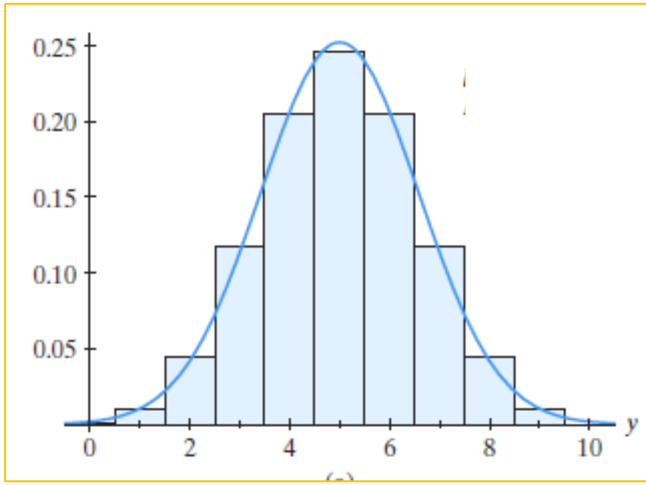
and variance $\sigma^2 = np(1-p)$.

$$\text{By } CLT, \ W = \frac{Y-\mu}{\sigma} = \frac{Y-np}{\sqrt{np(1-p)}} = \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \text{ is } N(0,1) \text{ as } n \to \infty.$$

For *sufficiently large n*, Y is approximately $N(np, np(1-p))$,

and probabilities for $b(n, p)$ can be approxiamted by $N(np, np(1-p))$.

The graph for ***pmf*** ⟷ The graph of ***Histogram*** — ***Approximately*** → The graph of ***Normal distribution***



➢ Binomial distribution (c.n.t.)

$$P(Y=k)=P(k-\frac{1}{2}<Y<k+\frac{1}{2})$$

$$f(k)=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

$$Y\sim N(np,np(1-p)) \text{ for large } n$$

$$P(k-\frac{1}{2}<Y<k+\frac{1}{2})=P(\frac{k-1/2-np}{\sqrt{np(1-p)}}<\frac{Y-np}{\sqrt{np(1-p)}}<\frac{k+1/2-np}{\sqrt{np(1-p)}})$$

$$=\Phi\left(\frac{k+1/2-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{k-1/2-np}{\sqrt{np(1-p)}}\right).$$

$\Phi(\cdot)$ is the cdf for $N(0,1)$

➢ Quiz

Assume $Y\sim b(10,0.5)$. Q: compute $P(3\le Y<6)$.

➢ Quiz

$$\text{Assume } Y \sim b(10, 0.5). \qquad Q: \text{compute } P(3 \le Y < 6).$$

Solution:

①By definition, $P(3 \le Y < 6) = \sum_{k=3}^{5} P(Y = k) = \sum_{k=3}^{5} f(k)$

②By CLT, $Y = \sum_{i=1}^{10} X_i$, $X_1, \ldots, X_{10}$ are i.i.d. from $b(1, 0.5)$.

Y approximately follows $N(np, np(1-p)) = N(5, 2.5)$

$$P(3 \le Y < 6) = \sum_{k=3}^{5} P(Y = k) = \sum_{k=3}^{5} P(k - \frac{1}{2} < Y < k + \frac{1}{2}) = P(2.5 < Y < 5.5)$$

$$= P(\frac{2.5-5}{\sqrt{2.5}} < \frac{Y-5}{\sqrt{2.5}} < \frac{5.5-5}{\sqrt{2.5}}) = \Phi(0.316) - \Phi(-1.581) = 0.6240 - 0.0570 = 0.5670$$

Chapter 5 Distributions of Functions of random variables

Section 5.8 Chebyshev's Inequality and convergence in probability

➢ Motivation

Given the mean and variance of a distribution, it's possible to have a rough estimate of probability of certain events [Some more evidence for why ***Sample mean*** $\bar{X}$ is a good estimate of mean]

Theorem 5.8-1 [Chebyshev's Inequality]

If the RV X has a mean μ and variance σ^2, then, for every $k \geq 1$,

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Proof.

Consider the discrete RV case. Let $f(x): \bar{S} \to [0,1]$ be the pmf of X.

Then $\sigma^2 = E\left[(X-\mu)^2\right] = \sum_{x\in\bar{S}}(x-\mu)^2 f(x) = \sum_{x\in A}(x-\mu)^2 f(x) + \sum_{x\in\bar{S}-A}(x-\mu)^2 f(x)$

where $A = \left\{x \in \bar{S} \,\|X-\mu| \geq k\sigma\right\}$.

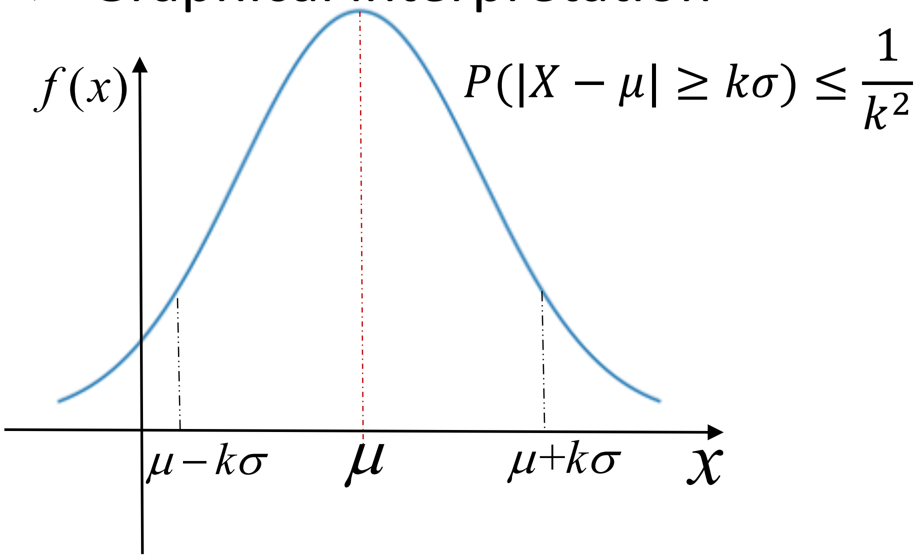
Since $\sum_{x\in\bar{S}-A}(x-\mu)^2 f(x) \geq 0$, we derive

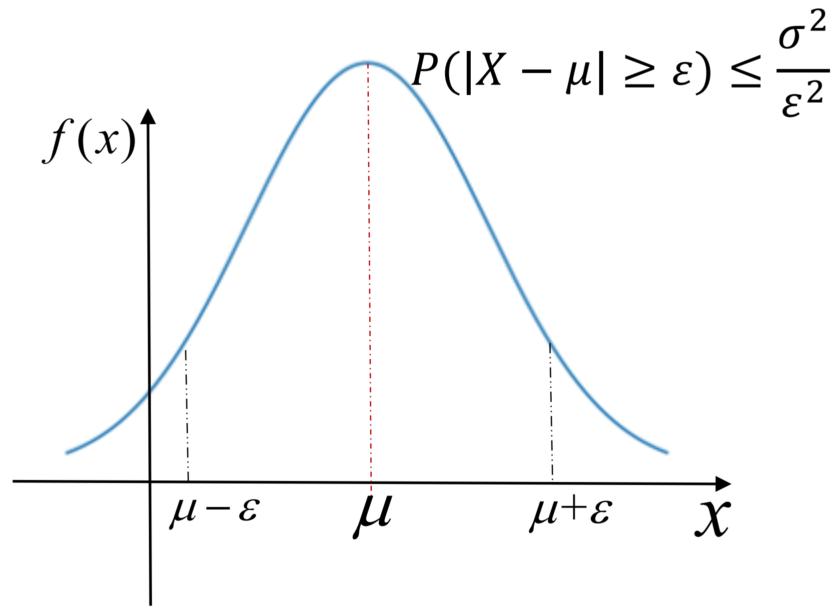
$$\sigma^2 \geq \sum_{x\in A}(x-\mu)^2 f(x) \geq k^2\sigma^2 \sum_{x\in A} f(x) = k^2\sigma^2 P(X \in A).$$

Corollary [Page 222]

$$\text{If } \varepsilon = k\sigma, \text{ then } P(|X-\mu| \ge \varepsilon) \le \frac{\sigma^2}{\varepsilon^2},$$

- Graphical interpretation





This links to the interpretation of σ^2, a measure of dispersion of X.

Example 1 [Page 222]

Note that the distribution of X is arbitrary!

Let X be a RV with mean 25 and varance 16.

Compute the loweer bound for $P(17 < X < 33)$ and upper bound for $P(|X-25| \ge 12)$.

Example 1 [Page 222] (c.n.t.)

Let X be a RV with mean 25 and varance 16.

Compute the loweer bound for $P(17 < X < 33)$ and upper bound for $P(|X-25| \geq 12)$.

Solution :

Lower bound for $P(17 < X < 33)$:

$$P(17 < X < 33) = P(|X-\mu| < 2\sigma) = 1 - P(|X-\mu| \geq 2\sigma) \geq 1 - \frac{1}{4}.$$

Upper bound for $P(|X-25| \geq 12) = P(|X-\mu| \geq 3\sigma) \leq \frac{1}{9}.$

Definition 5.8-1[Convergence in Probability]

A sequence of RVs $\{Y_n\}_{n=1}^{\infty}$ is said to connverge in probability to a constant μ, if for $\forall \varepsilon > 0$, $\lim_{n\to\infty} P(|Y_n - \mu| \geq \varepsilon) = 0$.

Theorem 5.8-2 [Law of Large Number]

Let $\overline{X}$ be the sample mean of a random sample $X_1, X_2, \ldots, X_n$ from a distribution with mean μ and finite variance σ^2, $\overline{X}$ ***converges in probability*** to μ.

In other words, $\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) = 0.$

Proof of theorem 5.8-2

Proof. Note that $E(\overline{X}) = \mu,\ Var(\overline{X}) = \frac{1}{n}\sigma^2.$

By corollary 5.8-1, for $\forall \varepsilon > 0$, we have

$$P(\left|\overline{X} - \mu\right| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \cdot \left(\frac{1}{n}\sigma^2\right)$$

Taking limits both sides yield:

$$0 \leq \lim_{n\to\infty} P(\left|\overline{X} - \mu\right| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{1}{\varepsilon^2} \cdot \left(\frac{1}{n}\sigma^2\right) = 0 \Rightarrow \lim_{n\to\infty} P(\left|\overline{X} - \mu\right| \geq \varepsilon) = 0.$$

or equivalently,

$$\lim_{n\to\infty} P(\left|\overline{X} - \mu\right| < \varepsilon) = 1.$$

Section 5.9 Limiting Moment Generating Functions

➢ Binomial distribution can be approximated by the Poisson distribution when n is large and p is fairly small.

Consistent with experiment

The *mgf* of $b(n,p)$ is $M(t)=(1-p+pe^t)^n$.

Let $np=\lambda$, we have

$$M(t)=(1-\frac{\lambda}{n}+\frac{\lambda}{n}e^t)^n$$

$$=\left[1+\frac{\lambda(e^t-1)}{n}\right]^n$$

Since $\lim_{n\to\infty}(1+\frac{b}{n})^n=e^b$, we have

$$\lim_{n\to\infty}M(t)=e^{\lambda(e^t-1)}$$

mgf for Poisson distribution

Example1 [P227]

Let $Y\sim b(50,1/25)$. Q: Compute $P(Y\le 1)$.

Example1 (c.n.t.)

Let $Y \sim b(50, 1/25)$. $\quad Q$: Compute $P(Y \leq 1)$.

Solution:

①By definition, $P(Y \leq 1) = P(Y = 0) + P(Y = 1) = \left(\frac{24}{25}\right)^{50} + 50\left(\frac{1}{25}\right)\left(\frac{24}{25}\right)^{2} = 0.4.$

②By approximation with Poisson distribution with $\lambda = np = 2$,

$$P(Y \leq 1) \approx \frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} = 3e^{-2} = 0.406.$$

Theorem 5.9-1 [convergence of mgfs]

Let $\{M_n(t)\}_{n=1}^{\infty}$ be a sequence of *mgfs* for t in an open interval around 0. If $\lim_{n\to\infty} M_n(t) = M(t)$, then the limit of the corresponding distributions must be the distribution corresponding to $M(t)$.

That is, convergence, for $|t| < h$, of mgfs to am mgf implies convergence of cdfs (thus implies the convergence of the distribution.)

- The proof of theorem 5.9-1 relies on the theory of Laplace transforms, which is omitted due to the beyond of the scope of this course.
- Then we show the proof of CLT by using theorem 5.9-1:

Theorem 5.6-1 (Central Limit Theorem))

If $\overline{X}$ is the sample mean of a random sample $X_1, X_2, \ldots, X_n$ of size n from a distribution with finite mean μ and finite positive variance σ^2, then the limit of the distribution of

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}}{\sigma}(\overline{X}-\mu) \sim N(0,1) \qquad \text{as } n \to \infty$$

equivalently, $\overline{X} \sim N(\mu, \frac{1}{n}\sigma^2)$ when $n \to \infty$. $\Leftrightarrow \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$ when $n \to \infty$.

Proof.

$$E(e^{tW}) = E\left\{\exp\left[\frac{t}{\sqrt{n}\sigma}\left(\sum_{i=1}^{n} X_i - n\mu\right)\right]\right\}$$

$$= E\left\{\exp\left[\frac{t}{\sqrt{n}}\cdot\frac{X_1-\mu}{\sigma}\right]\cdots\exp\left[\frac{t}{\sqrt{n}}\cdot\frac{X_n-\mu}{\sigma}\right]\right\}$$

$$= E\left\{\exp\left[\frac{t}{\sqrt{n}}\cdot\frac{X_1-\mu}{\sigma}\right]\right\}\cdots E\left\{\exp\left[\frac{t}{\sqrt{n}}\cdot\frac{X_n-\mu}{\sigma}\right]\right\}$$

Independence

Proof of theorem 5.6-1 (c.n.t.)

Let $m(t) = E\left\{\exp\left(t\dfrac{X_i - \mu}{\sigma}\right)\right\}$, $|t| < h$, be the common mgf for $Z_i = \dfrac{X_i - \mu}{\sigma}$, $i = 1, \ldots, n$.

Thus $E(e^{tw}) = \left[m\left(\dfrac{t}{\sqrt{n}}\right)\right]^n$.

Since $Z_i \sim N(0,1)$, $m(0) = 1$, $m'(0) = 0$, $m''(0) = 1$.

By using Taylor expansion, there exists $c \in [0, t]$ such that

$$m(t) = m(0) + m'(0)t + \frac{1}{2}m''(c)t^2 = 1 + \frac{1}{2}m''(c)t^2$$

$$= 1 + \frac{1}{2}t^2 + \frac{1}{2}t^2\left[m''(c) - 1\right].$$

$$E(e^{tw}) = \left[m\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left\{1 + \frac{1}{2}\frac{t^2}{n} + \frac{1}{2}\frac{t^2}{n}\left[m''(c) - 1\right]\right\}^n, \quad |t| < \sqrt{n}h. \qquad c \in \left[0, \frac{t}{\sqrt{n}}\right]$$

Since $m''(t)$ is continuous at $t = 0$ and $c \to 0$ as $n \to \infty$, we have

$$\lim_{n\to\infty} m''(c) - 1 = 1 - 1 = 0.$$

Proof of theorem 5.6-1 (c.n.t.)

Since $\lim_{n\to\infty}(1+\frac{b}{n})^n = e^b$, we have

$$\lim_{n\to\infty} E(e^{tw}) = \lim_{n\to\infty}\left\{1+\frac{1}{2}\frac{t^2}{n}+\frac{1}{2}\frac{t^2}{n}\left[m''(c)-1\right]\right\}^n$$

$$= \lim_{n\to\infty}\left\{1+\frac{1}{2}\frac{t^2}{n}\right\}^n = e^{t^2/2}.$$

mgf of *N(0,1)*

By theorem 5.9-1, $W = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.