Moment Generating Functions

Section 4.5 of text

Def. random variable X.

$$M_X(t) = E\left(e^{tX}\right)$$

disc. $$= \sum_x e^{tx} p(x)$$

contin. $$= \int_{\mathbb{R}} e^{tx} f(x)\,dx.$$

The MGF does not necessarily exist for any particular value of t.

(adv. Characteristic function does exist for every distribution)

Property A

If $M_X(t)$ exists for t in some open interval that contains 0, then it uniquely determines the probably distribution.

Property B

$$M_X^{(r)}(0) = E(X^r)$$

(rth derivative)

Take cts case

$$M_X(t) = E(e^{tX})$$

$$= \int_{\mathbb{R}} e^{tx} f(x)\,dx$$

$$= \int_{\mathbb{R}} \Bigg(\underbrace{\sum_{k=0}^{\infty} \frac{(tx)^k}{k!}}_{\text{Taylor exp.}} \Bigg) \cdot f(x)\,dx.$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_{\mathbb{R}} x^k f(x)\,dx.$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} E(X^k)$$

$$= 1 + \frac{t}{1!} \cdot E(X) + \frac{t^2}{2!} E(X^2)$$

$$+ \dots$$

$$\left. \frac{d^r}{dt^r} M_X(t) \right|_{t=0} = E(X^r)$$

A kth moment

$E(X^k)$

mean is first moment.

Poisson Example

$X \sim \text{Poisson}(\lambda)$

$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$

$M_X(t) = E(e^{tX})$

$$= \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k e^{-\lambda}}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!}$$

$$= e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

Expectation and variance from this.

$$\frac{d}{dt} e^{\lambda(e^t-1)} = e^{\lambda(e^t-1)} \cdot \lambda e^t$$

evaluate at $t=0$.

$$= e^{\lambda(1-1)} \cdot \lambda e^{0} = \lambda$$

Second moment

$$\frac{d}{dt} e^{\lambda(e^t-1)} \cdot \lambda e^{t}$$

$$= \lambda e^{t} e^{\lambda(e^t-1)} + \lambda e^{t} \cdot e^{\lambda(e^t-1)} \cdot \lambda$$

evaluating at $t=0$

$$\Rightarrow \lambda + \lambda^2$$

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2$$
$$= \lambda + \lambda^2 - \lambda^2$$
$$= \lambda$$

Property C

$Y = a + bX \quad a, b \in \mathbb{R}$

$$M_Y(t) = e^{at} M_X(bt)$$

Property D

Suppose $X \perp\!\!\!\perp Y$, def.

$Z = X + Y$

then $M_Z(t) = M_X(t) \cdot M_Y(t)$

<u>Pf.</u> $M_Z(t) = E(e^{tZ})$

$= E(e^{t(X+Y)})$

$= E(e^{tX} \cdot e^{tY})$

$= E(e^{tX})E(e^{tY})$ (by ind.)

$= M_X(t)\, M_Y(t).$

$\square$

Ex. Suppose $X \sim \text{pois}(\lambda)$

$Y \sim \text{pois}(\mu)$

$X \perp\!\!\!\perp Y$

what is distribution of $X+Y$

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

$$= e^{\lambda(e^t-1)} \cdot e^{\mu(e^t-1)}$$

$$= e^{(\lambda+\mu)(e^t-1)}$$

$\Rightarrow X+Y \sim \text{pois}(\lambda+\mu)$.

Assume $X_i \overset{ind}{\sim} \text{pois}(\lambda_i)$

then

$$\sum_{i=1}^{n} X_i \sim \text{pois}\left(\sum_{i=1}^{n} \lambda_i\right)$$

Pf. base case $n=2$.
just proved.

Induction step:
Assume holds for k

$$\sum_{i=1}^{k} X_i \sim \text{pois}\left(\sum_{i=1}^{k} \lambda_i\right)$$

call $\sum_{i=1}^{k} \lambda_i = \mu$.

call $\sum_{i=1}^{k} X_i = Y$

Take $Y + X_{k+1} \sim \text{pois}(\mu + \lambda_{k+1})$

$$= \sum_{i=1}^{k+1} X_i \sim \text{pois}\left(\sum_{i=1}^{k+1} \lambda_i\right)$$

$\square$

$\boxed{\text{Logic of MLE}}$

Suppose we have a dataset that we observe

$$X_1, \dots, X_n \overset{iid}{\sim} p_\theta(x).$$

θ unknown.

The joint density is

$$\prod_{i=1}^{n} p_\theta(x_i) := L_n(\theta; \underline{x})$$

we don't work products

The log likelihood is

$$\log\left(\prod_{i=1}^{n} p_\theta(x_i)\right) = \ell_n(\theta; \underline{x})$$

$$\sum_{i=1}^{n} \log p_\theta(x_i)$$

Poisson example

$X_1, \dots, X_n$, $X_i \overset{iid}{\sim} \text{pois}(\lambda)$

λ unknown. $P_\lambda(X_i = k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$

$$P_\lambda(X_1, \dots, X_n) = L(\lambda)$$

$$= \prod_{i=1}^{n} P_\lambda(X_i = x_i)$$

$$\Rightarrow \ell(\lambda) = \log\left(\prod_{i=1}^{n} P_\lambda(X_i = x_i)\right)$$

$$= \sum_{i=1}^{n} \log P_\lambda(X_i = x_i)$$

$$= \sum_{i=1}^{n} \log\left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)$$

$$= \sum_{i=1}^{n}\left(x_i \log\lambda - \lambda - \log(x_i!)\right)$$

$$= \log \lambda \sum_{i=1}^{n} x_i - n\lambda - \sum_{i=1}^{n} \log(x_i!))$$

$$= \ell_n(\lambda; \underline{x})$$

$$\frac{\partial \ell_n(\lambda; \underline{x})}{\partial \lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$$

$$\frac{1}{n} \sum_{i=1}^{n} x_i = \hat{\lambda} = \bar{x}$$

Section 4.7. Pb 54

X, Y, Z are uncorrelated r.v.

have variances $\sigma^2_X, \sigma^2_Y, \sigma^2_Z$

Def. $U = Z + X$

$W = Z + Y$

$\mathrm{cov}(U, W)$, ρ_{UW}

$\mathrm{cov}(U, W) = \mathrm{cov}(Z+X, Z+Y)$

$$= \mathrm{cov}(Z+X, Z) + \mathrm{cov}(Z+X, Y)$$

$$= \mathrm{Cov}(Z, Z) + \mathrm{cov}(X, Z) + \mathrm{cov}(Z, Y) + \mathrm{cov}(X, Y)$$

$$= \sigma_Z^2$$

$$\mathrm{Cov}(X, Y) = E\big((X - E(X))(Y - E(Y))\big)$$

if

$$X = Y \Rightarrow E\big((X - E(X))^2\big)$$

$$\rho_{UW} = \frac{\mathrm{cov}(U, W)}{\sigma_U \cdot \sigma_W}$$

$$\sigma_v^2 = var(z+x)$$

$$= var(z) + var(x) + \cancel{2cov(z,x)}^{\,0}$$

$$\sigma_v = \sqrt{var(z) + var(x)}$$

Sec 4.7. QSS

Let $T = \sum_{k=1}^{n} k X_k$

$X_k \overset{iid}{\sim} [\mu, \sigma^2]$

ET, $VarT$

$$ET = \sum_{k=1}^{n} k\, EX_k = \sum_{k=1}^{n} k\mu$$

$$= \frac{n(n+1)}{2}\mu.$$

$$\mathrm{Var}\,T = \sum_{k=1}^{n} k^2\, \mathrm{Var}(X_k)$$

$$= \sigma^2 \sum_{k=1}^{n} k^2 = \sigma^2 \frac{n(n-1)(2n+1)}{6}$$

Section 5.4 Q4

N is poisson. $E(N)=100$.

$N \sim$ poisson(100).

Use normal approx to find

Δ such that

$$P(100-\Delta < N < 100+\Delta) \approx 0.9$$

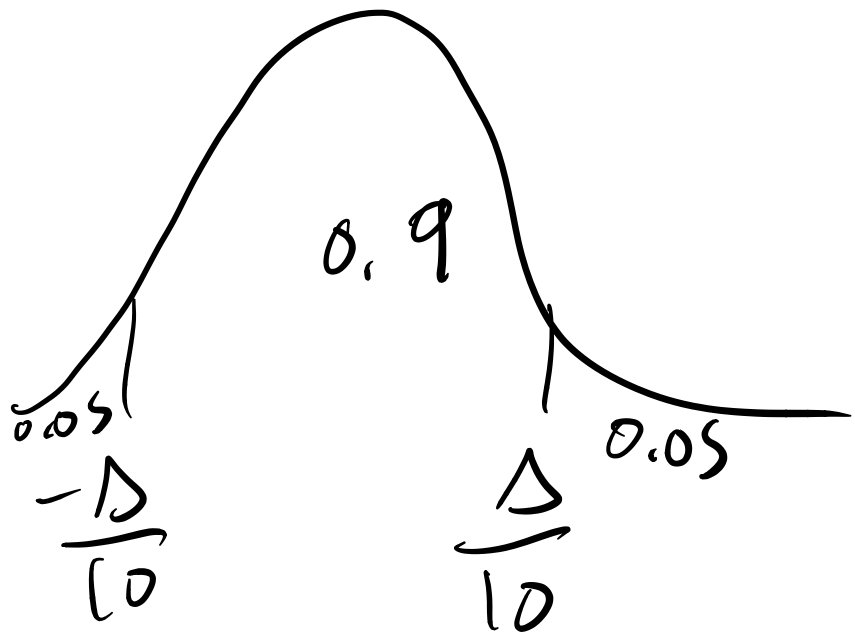
$$Var(N) = 100.$$

$$P(100-\Delta < N < 100+\Delta)$$

$$= P\left(\frac{100-\Delta-100}{10} < \frac{N-100}{10} < \frac{100+\Delta-100}{10}\right)$$

$$= P\left(\frac{-\Delta}{10} < z < \frac{\Delta}{10}\right)$$

$$= 0.9$$



$z = 1.64$

$$1.64 = \frac{\Delta}{10} \Rightarrow \Delta = 16.4$$

Section 5.4 Q5

Using MGF's show that

as $n \to \infty$, $p \to 0$

and $np \to \lambda$

$\text{bin}(n,p) \to \text{pois}(\lambda)$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$M_X(t) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k \cdot (1-p)^{n-k}$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

$$= \left(pe^t + (1-p)\right)^n$$

$$= \left(1 + (e^t - 1)p\right)^n$$

$$\lambda = np. \Rightarrow p = \frac{\lambda}{n}$$

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

now we take lim as $n \to \infty$

$$\lim_{n \to \infty} \left(1 + \frac{\lambda(e^t - 1)}{n}\right)^n$$

$$= e^{\lambda(e^t - 1)}$$

Section 5.4 Q16

$X_1, \ldots, X_{20}$ are ind r.v.

w/ density $f(x) = 2x$

$0 \le x \le 1$

Def. $S = X_1 + \dots + X_{20}$

Approximate $P(S \le 10)$

$$P\left(\frac{S - E(S)}{sd(S)} \le \frac{10 - E(S)}{sd(S)} \right)$$

$\sim N(0,1)$

Section 5.4 Q17

measurements mean μ, variance $\sigma^2 = 25$.

How large n should be so that

$$P(|\bar{X} - \mu| < 1) = 0.95.$$

$$P(-1 < \bar{X} - \mu < 1)$$

$$E(\bar{X}) = E\frac{1}{n}\sum X_i$$

$$= \frac{1}{n}\sum EX_i = \frac{1}{n}\cdot n\mu$$

$$= \mu.$$

$$Var(\bar{X}) = Var\,\frac{1}{n}\sum X_i$$

$$= \frac{1}{n^2}\sum Var(X_i)$$

$$= \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}.$$

$$sd(\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{s}{\sqrt{n}}$$

$$P(-1 < \bar{X} - \mu < 1)$$

$$= P\left(\frac{\sqrt{n}(-1)}{5} < \frac{\sqrt{n}(\bar{X} - \mu)}{5} < \frac{\sqrt{n}}{5}\right)$$

$$= 0.95$$

$$N(0,1)$$

$$\frac{\sqrt{n}}{5} = 1.96$$

$$n = 97$$

Sec 5.4 Q18

$X_i \overset{ind.}{\sim} [15, 10^2]$

$P\left(\sum_{i=1}^{100} X_i > 1700 \right)$

$S_{100} = \sum_{i=1}^{100} X_i \quad E(S_{100}) = 100 \cdot 15$

$= 1500$

$Var(S_{100}) = 100 \cdot 100.$

$P(S_n > 1700)$

$= P\left(\frac{S_n - 1500}{100} > \frac{1700-1500}{100} \right)$

$$= 1 - P(Z \le 2)$$

$$= 1 - \Phi(2)$$

$$= 0.0228$$

$$Var\left(\sum_{i=1}^{100} X_i\right)$$

$$= \sum_{i=1}^{100} Var(X_i) = \sum_{i=1}^{100} \sigma^2$$

$$= 100 \cdot \sigma^2$$