Chebyshev inequality

Assume $X \in \mathbb{R}$ is a random variable with finite $E(X)$ and $Var(X)$. Set $\varepsilon > 0$.

1) Define $Z(x) := X - E(X)$

2) Define $Y(x) = \begin{cases} 0 & \text{if } |Z(x)| < \varepsilon \\ \varepsilon^2 & \text{if } |Z(x)| \geq \varepsilon \end{cases}$

3) From this follows: $Y(x) \leq |Z(x)|^2$

<u>Check:</u> If $|Z| < \varepsilon$, then $Y = 0$ ✓
If $|Z| \geq \varepsilon$, then $Y = \varepsilon^2$ ✓

4) Note that $Y = \varepsilon^2$ implies $|Z| \geq \varepsilon$

<u>Proof:</u> $Var(X) = E(|X - E(x)|^2) = E(|Z|^2)$

$$\overset{3)}{\geq} E(Y) = 0 \cdot P(Y=0) + \varepsilon^2 \cdot P(Y = \varepsilon^2)$$

$$\overset{4)}{=} \varepsilon^2 P(|Z| \geq \varepsilon)$$

$$\Rightarrow P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2} \qquad \square$$

Law of large numbers

Assume $X_1, \dots, X_n$, $n \in \mathbb{N}$, are independent random variables with equal expectation values and finite variances $Var(X_i) \leq M < \infty$ $\forall i \in [1, n]$.

1) Set $X = \frac{1}{n} \sum_{i=1}^{n} X_i \Rightarrow E(X) = E(X_i) = \mu$

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

we use that $Var(\lambda X) = \lambda^2 Var(X)$ for scalars λ, and

$Var(X+Y) = Var(X) + Var(Y)$ for independent random variables X, Y

3) Apply Chebyshev: $P(|X - \underset{\overset{\shortparallel}{\mu}}{E(X)}| \geq \varepsilon)$

$$\leq \frac{Var(X)}{\varepsilon^2} = \frac{M}{n\varepsilon^2} \quad \square$$

Generating moments

Let $\varphi_X(t)$ be the moment-generating function of random variable X with probability distribution $P(x)$.
The n^{th} moment is obtained by taking the n^{th} derivative and setting $t=0$:

$$\frac{d^n}{dt^n}\varphi_X(t)\Big|_{t=0} = \sum_{i=0}^{\infty}\frac{1}{i!}\int_{-\infty}^{\infty} dx\, \underbrace{\frac{d^n}{dt^n}(tx)^i}\, P(x)$$

all terms $i<n$: vanish from differentiation

all terms $i>n$: t^{i-n} remains, vanishes for $t=0$

$\underline{i=n}$: t gone, and we gain a factor $n!$ from differentiating n times, which cancels $\frac{1}{n!}$

$$\Longrightarrow \frac{d^n}{dt^n}\varphi_X(t) = \int_{-\infty}^{\infty} dx\, x^n P(x) = \langle x^n \rangle \quad \square$$

Moment-generating fct. of Gaussian

Gauss: $$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\Rightarrow E(e^{tx}) = \frac{1}{\sqrt{2\pi}\sigma}\int dx\, e^{\underbrace{-\frac{(x-\mu)^2}{2\sigma^2}+tx}_{=B}}$$

Look at exponent:

$$B = \frac{1}{2\sigma^2}\left(-x^2+2x\mu-\mu^2+2\sigma^2 tx\right)$$

$$= -\frac{1}{2\sigma^2}\left(x^2-2x(\mu+t\sigma^2)+\mu^2\right)$$

complete square

$$= -\frac{1}{2\sigma^2}\left(x^2-2x(\mu+t\sigma^2)+(\mu+t\sigma^2)^2-(\mu+t\sigma^2)^2+\mu^2\right)$$

$$= -\frac{1}{2\sigma^2}\left((x-\mu-t\sigma^2)^2-(\mu+t\sigma^2)^2+\mu^2\right)$$

Insert back: $$E(e^{tx}) = e^{\underbrace{\frac{(\mu+t\sigma^2)^2-\mu^2}{2\sigma^2}}_{=C}} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\sigma}\int dx\, e^{-\frac{(x-\mu+t\sigma^2)^2}{2\sigma^2}}}_{=1}$$

(normal Gauss distr.)

$$C = \frac{\mu^2+2t\sigma^2\mu+t^2\sigma^4-\mu^2}{2\sigma^2}$$

$$= t\mu+\frac{1}{2}t^2\sigma^2$$

$$\Rightarrow \boxed{E(e^{tx}) = e^{t\mu+\frac{1}{2}t^2\sigma^2}} \quad \square$$

Central limit theorem (CLT)

Let $Z = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$, where X_i are indep. random variables with finite expectation value and variance, $E(X_i) \leq A < \infty$ $\forall i$

$\mathrm{Var}(X_i) \leq B < \infty$

To prove the CLT, we only need three properties:

1) If the moment-gen. fct. exists, then all cumulants are finite.

2) Cumulants are additive.

<u>Proof:</u> $\ln(E(e^{tZ})) = \ln(E(\prod_{i=1}^{n} e^{\frac{tX_i}{\sqrt{n}}}))$

$= \ln(\prod_{i=1}^{n} E(e^{\frac{tX_i}{\sqrt{n}}})) = \sum_{i=1}^{n} \ln E(e^{\frac{tX_i}{\sqrt{n}}})$

X_i indep. $\nearrow$

The j^{th} cumulant is then given by $k_j(Z) = \frac{d^j}{dt^j} \ln(E(e^{tZ}))\Big|_{t=0}$

$= \sum_{i=1}^{n} \underbrace{\frac{d^j}{dt^j} \ln(E(e^{\frac{tX_i}{\sqrt{n}}}))\Big|_{t=0}}$

$= \sum_{i=1}^{n} k_j(\frac{X_i}{\sqrt{n}})$ $\square$

3) $\kappa_j(z) = n^{-j/2} \sum_{i=1}^{n} \kappa_j(x_i)$

<u>Proof:</u> We show that $\kappa_j(az) = a^j \kappa_j(z)$ for a scalar a.

$$\kappa_j(az) = \frac{d^j}{dt^j} \ln E(e^{taz})\Big|_{t=0}$$

substitute $t' = t \cdot a$

$$\Rightarrow \frac{d}{dt} = \frac{dt'}{dt}\frac{d}{dt'} = a\frac{d}{dt'} \qquad \stackrel{\downarrow}{=} a^j \frac{d^j}{dt'^j} \ln E(e^{t'z})\Big|_{t'=0}$$

$$\Rightarrow \frac{d^j}{dt^j} = a^j \frac{d^j}{dt'^j} \qquad = a^j \kappa_j(z)$$

<u>Proof of CLT:</u>

1) $\kappa_1(z) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \kappa_1(x_i) \leq \frac{1}{\sqrt{n}} \cdot n \cdot A = \sqrt{n}\, A$

2) $\kappa_2(z) = \frac{1}{n} \sum_{i=1}^{n} \kappa_2(x_i) \leq \frac{1}{n} \cdot n \cdot B = B$

3) For κ_j with $j > 0$: suppressed by factor $n^{-j/2+1}$, i.e., for $n \to \infty$, all but the first two cumulants disappear

$\Rightarrow z$ has $\varphi_z(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$ with $\mu = \kappa_1(z)$, $\sigma^2 = \kappa_2(z)$

$\Rightarrow P(z)$ is a Gaussian (obtained via inv. Lapl. trans.)

Markov inequality

Show: $P(X \geq \varepsilon) \leq \dfrac{E(X)}{\varepsilon} \quad \forall \varepsilon > 0, X \geq 0$

Proof: $E(X) = \int x \, p(x) \, dx$

$$= \underbrace{\int_0^{\varepsilon} x \, p(x) \, dx}_{\geq 0} + \int_{\varepsilon}^{\infty} x \, p(x) \, dx$$

$$\geq \int_{\varepsilon}^{\infty} x \, p(x) \, dx \geq \varepsilon \underbrace{\int_{\varepsilon}^{\infty} p(x) \, dx}_{= P(X \geq \varepsilon)}$$

$$\Rightarrow E(X) \geq \varepsilon \, P(X \geq \varepsilon) \quad \square$$

Chernoff bounds

Random variable X, gen. function $\varphi_X(t) = E(e^{tX})$
Assume $|\varphi_X(t)| < \infty$ for $|t| < b$ and $b > 0$.

→ **Markov:** $P(X \geq \varepsilon) \leq \dfrac{E(X)}{\varepsilon}$

scale by $t > 0$: $P(X \geq \varepsilon) = P(tX \geq t\varepsilon) \leq \dfrac{E(tX)}{t\varepsilon}$

apply mon. incr. transformation $P(tX \geq t\varepsilon) = P(e^{tX} \geq e^{t\varepsilon}) \leq \dfrac{E(e^{tX})}{e^{t\varepsilon}}$

$= e^{-t\varepsilon}\varphi_X(t)$

apply inf $\Rightarrow$ $P(\underbrace{X \geq \varepsilon}) \leq \inf\limits_{0 \leq t \leq b} e^{-t\varepsilon}\varphi_X(t)$ $\square$

note: one-sided bound! (no absolute value taken)

bound holds for any t, so we choose the t that produces the tightest bound

Basically: Chernoff bound = applying Markov bound to e^{tX}

Moment-gen fct. for scaled sum of Gaussians

For a single Gaussian RV, the moment-gen. fct. is given by $\varphi_X(t) = e^{\mu t + \frac{t^2\sigma^2}{2}}$

μ: mean, σ^2: variance

Let $\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $z = \bar{x}_n - E(\bar{x}_n)$

We know the following:

1) A sum of Gaussian RV is itself a Gaussian RV.

2) $E(z) = E(\bar{x}_n) - E(\bar{x}_n) = 0$

3) $\mathrm{Var}(z) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\underbrace{\sum \mathrm{Var}(x_i)}_{=n\sigma^2}$

$$= \frac{\sigma^2}{n}$$

Thus, z is a Gaussian RV with mean 0 and variance $\frac{\sigma^2}{n}$, and its moment-gen. fct. is $\varphi_z(t) = e^{\frac{t^2\sigma^2}{2n}}$.

Chernoff bound for sum of Gaussian RVs

Chernoff: $P(Z \geq \varepsilon) \leq \inf_t \left[e^{-\varepsilon t} \varphi_Z(t) \right]$

and $\varphi_Z(t) = e^{\frac{t^2 \sigma^2}{2n}}$

1) Find minimum: $\frac{\partial}{\partial t} e^{\frac{t^2 \sigma^2}{2n} - \varepsilon t} \overset{!}{=} 0$

$$\iff e^{\frac{t^2\sigma^2}{2n} - \varepsilon t} \left[\frac{t\sigma^2}{n} - \varepsilon \right] = 0$$

$$\iff t = \frac{\varepsilon n}{\sigma^2}$$

2) Plug in: $P(Z \geq \varepsilon) \leq e^{-\frac{\varepsilon^2 n}{\sigma^2}} \cdot e^{\frac{\varepsilon^2 n}{2\sigma^2}}$

$$= e^{-\frac{n\varepsilon^2}{2\sigma^2}}$$

3) Due to symmetry: $P(Z \leq -\varepsilon) \leq e^{-\frac{n\varepsilon^2}{2\sigma^2}}$

$$\implies P(|Z| \geq \varepsilon) \leq 2e^{-\frac{n\varepsilon^2}{2\sigma^2}}$$

Popoviciu's inequality

Show: $Var(X) \leq \frac{(b-a)^2}{4}$ if $E(X) = \mu < \infty$ and $P(a \leq X \leq b) = 1$

1) $E(\underbrace{(b-x)}_{\geq 0}\underbrace{(x-a)}_{\geq 0}) = \mu(b+a) - ab - E(x^2) \geq 0$

2) $$\begin{aligned} Var(x) &= E(x^2) - \mu^2 \\ &\leq \mu(a+b) - ab - \mu^2 \\ &= (b-\mu)(\mu-a) \end{aligned}$$

3) Note that $a \cdot b \leq \left(\frac{a+b}{2}\right)^2$ or $\sqrt{ab} \leq \frac{a+b}{2}$ if $a, b \geq 0$

Why? $\rightsquigarrow$ $$\begin{aligned} \left(\frac{a+b}{2}\right)^2 - ab &= \frac{1}{4}(a^2+2ab+b^2) - ab \\ &= \frac{1}{4}(a^2-2ab+b^2) \\ &= \left(\frac{a-b}{2}\right)^2 \geq 0 \end{aligned}$$

$$\Longrightarrow \left(\frac{a+b}{2}\right)^2 \geq a \cdot b \quad \square$$

4) Apply this: $Var(X) \leq \left(\frac{b-\mu+\mu-a}{2}\right)^2 = \frac{(b-a)^2}{4} \quad \square$