Lecture Stat 302
Introduction to Probability - Slides 15

AD

March 2010

Continuous Random Variable

- Let X a (real-valued) continuous r.v.. It is characterized by its pdf $f : \mathbb{R} \rightarrow [0, \infty)$ which such that for any set A of real numbers

$$P(X \in A) = \int_A f(x)\, dx.$$

and its distribution function

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(y)\, dy.$$

- For any real-valued function $g : \mathbb{R} \rightarrow \mathbb{R}$, we have

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)\ f(x)\, dx$$

Normal Random Variables

- Also known as Gaussian random variables in the literature.
- We say that X is a normal r.v. of parameters $\left(\mu, \sigma^2\right)$ if its pdf is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

- The normal distribution is often used to describe, at least approximately, any variable that tends to cluster around the mean; e.g. the heights of USA males are roughly normally distributed. A histogram of male heights will appear similar to a bell curve, with the correspondence becoming closer if more data are used.

Properties of Normal Random Variables

- It can indeed be checked that

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \sqrt{2\pi}\sigma.$$

- We have also

$$E(X) = \mu$$

  and

$$Var(X) = \sigma^2$$

- Hence μ is referred to as the mean and σ^2 as the variance.

The 68-95-99.7 Rule

- We have

$$
\begin{aligned}
P(\mu - \sigma \leq X \leq \mu + \sigma) &\approx 0.68, \\
P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &\approx 0.95, \\
P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &\approx 0.997.
\end{aligned}
$$

- This helps doing quickly some approximate calculations.
- The distribution of the scores of the more than 1.3 million high school seniors in 2002 who took the SAT verbal exam is close to normal with $(\mu, \sigma^2) = (504, 111^2)$.
- Hence 95% of the SAT scores are between $504 - 222 = 282$ and $504 + 222 = 276$. The other 5% of scores lie outside this range.

Properties of Normal Random Variables

- Let X be a normal r.v. of parameters (μ, σ^2) and consider the new r.v. Y such that
$$Y = aX + b$$
then we know that
$$\begin{aligned} E(Y) &= aE(X) + b = a\mu + b, \\ Var(Y) &= a^2 Var(X) = a^2\sigma^2. \end{aligned}$$
- A much stronger result is true, Y is a normal r.v. of parameters $(a\mu + b, a^2\sigma^2)$.

Properties of Normal Random Variables

- For $a > 0$, we have
$$P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$
- The *chain rule* tells us that $[u(v(y))]' = v'(y) \cdot u'(v(y))$ so
$$\begin{aligned} f_Y(y) &= \tfrac{1}{a} f_X\left(\tfrac{y-b}{a}\right) \\ &= \tfrac{1}{a} \tfrac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma a} \exp\left(-\frac{(y-b-a\mu)^2}{2\sigma^2 a^2}\right) \end{aligned}$$
- For $a < 0$, we use
$$P(Y \leq y) = P(aX + b \leq y) = P\left(X \geq \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right)$$
and
$$f_Y(y) = \frac{-1}{a} f_X\left(\frac{y-b}{a}\right) = \frac{1}{\sqrt{2\pi}\sigma\,|a|} \exp\left(-\frac{(y-b-a\mu)^2}{2\sigma^2 a^2}\right)$$

Cumulative Distribution Function

- Consider X a normal r.v. of parameters $\left(\mu = 0, \sigma^2 = 1\right)$; known as *standard* r.v. in the literature.
- It is customary to denote $\Phi\left(x\right)$ the cdf of X; i.e.

$$\Phi\left(x\right) = P\left(X \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy.$$

- $\Phi\left(x\right)$ does not admit an analytical expression but is tabulated for $x \geq 0$.
- One can easily show that

$$\Phi\left(-x\right) = P\left(X \leq -x\right) = P\left(X \geq x\right) = 1 - \Phi\left(x\right)$$

Standardizing normal variables

- Let X a normal r.v. of mean μ and variance σ^2.
- Define the new r.v.

$$Z = \frac{X - \mu}{\sigma}$$

  then Z is a standard normal r.v.
- Hence

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

Example

- Let X a normal r.v. of mean $\mu = 2$ and variance $\sigma^2 = 25$. Assume you want to compute using the table of $\Phi(x)$ (a) $P(1 \leq X \leq 4)$, (b) $P(X > 0)$ and (c) $P\left((X-2)^2 > 5\right)$
- (a) We have

$$\begin{aligned} P(1 \leq X \leq 4) &= P\left(\frac{1-2}{5} \leq \frac{X-2}{5} \leq \frac{4-2}{5}\right) \\ &= P\left(\frac{-1}{5} \leq Z \leq \frac{2}{5}\right) = \Phi\left(\frac{2}{5}\right) - \Phi\left(-\frac{1}{5}\right) \\ &= \Phi\left(\frac{2}{5}\right) - \left(1 - \Phi\left(\frac{1}{5}\right)\right) \end{aligned}$$

where Z is a normal r.v. of mean 0 and variance 1; i.e. a standard normal r.v.

Example

- (b) We have

$$\begin{aligned} P(X>0) &= P\left(\frac{X-2}{5} > \frac{-2}{5}\right) = P\left(Z > \frac{-2}{5}\right) \\ &= 1-\Phi\left(-\frac{2}{5}\right) = \Phi\left(\frac{2}{5}\right) \end{aligned}$$

- (c) We have

$$\begin{aligned} P\left((X-2)^2 > 5\right) &= P\left(\frac{(X-2)^2}{25} > \frac{1}{5}\right) = P\left(Z^2 > \frac{1}{5}\right) \\ &= P\left(Z > \frac{1}{\sqrt{5}}\right) + P\left(Z < -\frac{1}{\sqrt{5}}\right) \\ &= 1-\Phi\left(\frac{1}{\sqrt{5}}\right) + \Phi\left(-\frac{1}{\sqrt{5}}\right) \\ &= 2\left(1-\Phi\left(\frac{1}{\sqrt{5}}\right)\right) \end{aligned}$$

Example: Signal Transmission

- A binary message - either 0 or 1 - is transmitted through the atmosphere from A to B. The value 2 is sent when the message is 1 and the value -2 is sent when the message is 0. At the location B of the receiver, the message received is corrupted by some channel noise; that is if the signal $X = x$ has been transmitted then at the receiver we observe

$$R = x + N$$

  where the noise is assumed to be a standard normal r.v.
- At the receiver, the following decoding scheme is used. If $R \geq 0.5$ then we conclude that 1 has been transmitted. If $R < 0.5$ then we conclude that 0 has been transmitted.
- What is the probability of decoding correctly the transmitted message when we transmit 0 and when we transmit 1?

Example: Signal Transmission

- If we transmit 0, then $R = -2 + N$ is an normal r.v. of mean -2 and variance 1 so

$$\begin{aligned} P(R < 0.5) &= P\left(\frac{R+2}{1} < \frac{0.5+2}{1}\right) \\ &= P(Z < 2.5) = \Phi(2.5) \approx 0.999 \end{aligned}$$

- If we transmit 1, then $R = 2 + N$ is an normal r.v. of mean 2 and variance 1 so

$$\begin{aligned} P(R > 0.5) &= P\left(\frac{R-2}{1} > \frac{0.5-2}{1}\right) \\ &= P(Z > -1.5) = \Phi(1.5) \approx 0.933 \end{aligned}$$

- Generalization of this idea = Viterbi algorithm.

Normal Approximation to the Binomial Distribution

- Consider X a binomial r.v. of parameters n, p then we know that

$$E(X) = np, \; Var(X) = np(1-p).$$

- We have already seen that it is possible to approximate X by a Poisson distribution of parameter $\lambda = np$.
- As $np \to \infty$, it can be shown that X can be approximated by a normal r.v. with $\mu = np$ and $\sigma^2 = np(1-p)$ so

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) \\ &\approx \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

Example: Bald men

- If 10% of men are bald, what is the probability that fewer than 100 in a random sample of 818 men are bald?
- Let X be the number of bald men in a random sample of 818 men, this is a Bernoulli r.v. of parameters $p = 0.1$ and $n = 818$.
- We are interested in computing $P(X \leq 100)$. We can use the standard binomial but this is tiedous. We use the normal approximation where

$$\mu = np = 81.8, \ \sigma = \sqrt{np(1-p)} = 8.5802$$

so

$$\begin{aligned} P(0 \leq X \leq 100) &= \Phi\left(\frac{100 - 81.8}{8.5802}\right) - \Phi\left(\frac{-81.8}{8.5802}\right) \\ &\approx 0.983. \end{aligned}$$

Example: Threshold signal

- Assume to transmit a random signal X which follows a normal distribution (μ, σ^2). The receiver only detects signals above a given threshold m so that what is observed is

$$Y = \begin{cases} X & \text{if } X \geq m \\ 0 & \text{if } X < m \end{cases}$$

- Compute the expected value of the received signal Y?

Example: Threshold signal

- We have

$$E(Y) = \int_m^\infty x \, \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$
$$= \int_m^\infty (x-\mu) \, \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx + \mu \int_m^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

where we use $x = (x-\mu) + \mu$.

- Now we have $\int_m^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1 - \Phi\left(\frac{m-\mu}{\sigma}\right)$ and

$$\begin{aligned}
&\int_m^\infty (x-\mu) \, \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\
&= \left[\frac{-\sigma^2}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right]_m^\infty \\
&= \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right)
\end{aligned}$$

so

$$E(Y) = \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) + \mu \left(1 - \Phi\left(\frac{m-\mu}{\sigma}\right)\right).$$

Exercise: Stein's identity

- Let X a normal random variable of mean μ and variance σ^2 then show

$$E\left[(X-\mu)\,g(X)\right]=\sigma^2 E\left[g'(X)\right]$$

  when both sides exist.
- We have

$$E\left[(X-\mu)\,g(X)\right]=\int_{-\infty}^{\infty} g(x)\;\;(x-\mu)\;\;\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\;dx$$

  so by integration by parts

$$\begin{aligned}E\left[(X-\mu)\,g(X)\right] \;&=\; \left[g(x)\times\frac{-\sigma^2}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right]_{-\infty}^{\infty}\\ &\quad+\int_{-\infty}^{\infty} g'(x)\;\;\times\frac{\sigma^2}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx\end{aligned}$$